UNIVERSITÀ DEGLI STUDI DI BERGAMO
Department of Economics

# WORKING PAPERS

## Measuring salience in macroeconomic time series

Francesca Papagni, Emilio Zanetti Chini

March 2026 - WP N. 40 Year 2026

**Working papers – Department of Economics**
**n. 40**

# Measuring salience in macroeconomic time series




**Francesca Papagni, Emilio Zanetti Chini**

**Università degli Studi di Bergamo**
**2026**

Measuring salience in macroeconomic time series / Francesca Papagni, Emilio Zanetti Chini.
Bergamo: Università degli Studi di Bergamo, 2026.
Working papers of Department of Economics, n. 40
ISSN: 2974-5586
DOI: 10.13122/WPEconomics_40

Il working paper è realizzato e rilasciato con licenza
Attribution Non commercial license (CC BY-NC 4.0)
https://creativecommons.org/licenses/by-nc/4.0/
La licenza prevede la possibilità di ridistribuire liberamente l'opera, a patto che venga citato il nome degli autori e che la distribuzione dei lavori derivati non abbia scopi commerciali.




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/10446/323865

# Measuring salience in macroeconomic time series

Francesca Papagni *[1] and Emilio Zanetti Chini[1]

[1]Department of Economics, University of Bergamo

March 17, 2026

Preliminary version.

**Abstract**

This paper proposes a statistical framework to detect salience and its effects in time series data. We focus on two fundamental dimensions of salience—*surprise* and *prominence*—and translate them into well-defined stochastic features. Surprise arises from localized and unexpected deviations from regular behavior, while prominence is associated with persistence and long-lasting patterns generated by strong temporal dependence. We show that these dimensions affect the spectral representation of a stationary process by introducing specific features. A general spectral model based on a power transformation of the spectral density allows us to interpret the selection of the power parameter from a weighted Whittle likelihood perspective and to detect salience-driven attention and its implications on expectations. Simulation evidence illustrates how the proposed approach discriminates between non-salient dynamics and salient time series behavior. A real-data application on inflation and consumer survey expectations shows how to identify salience-induced biases in inflation forecasts.

*Key words*: Salience, Spectral Methods, Time Series Modeling, Inflation Expectations, Behavioral Macroeconomics, Forecasting

*Corresponding author. Address: Department of Economics, University of Bergamo, Via dei Caniana 2, 24127, Bergamo, Italy; E-mail: francesca.papagni@unibg.it.

# 1 Introduction

Salience plays a central role in theories of attention and economic choice. In the framework developed by Bordalo et al. (2022), salient features of the environment attract disproportionate attention and therefore affect judgments, beliefs, and decisions. While this idea has proved highly influential in behavioral economics, its empirical implementation for historical macroeconomic data remains underdeveloped. In particular, there is still limited evidence on how salience can be identified and measured in time series settings, where observations are not independent and where attention may be shaped by the temporal structure of the data.

We propose a statistical framework to identify salient characteristics in time series and detect possible subsequent effects on shaping individuals' expectations through the attention mechanism. We focus on two dimensions emphasized in the salience literature, namely *surprise* and *prominence*, and translate them into features of a stochastic process. Surprise is associated with departures from regularity that stand out relative to recent behavior, while prominence refers to highly available features of the time series, and is therefore linked to components of the process that are systematically more visible, thus contributing disproportionately to overall variation. In a time series context, these dimensions are embedded in the temporal dependence structure of the process and can be characterized in the spectral domain. Specifically, salient features attributable to surprise and prominence translate into spectral peaks at certain frequencies.

Our approach is based on the spectral representation of stationary processes. We refer to a class of transformed spectral models indexed by a power parameter $p$ (Proietti and Luati, 2015), where this parameter can be interpreted as capturing the extent to which emphasizing certain frequency components improves the spectral model fit in terms of a weighted Whittle likelihood criterion. A selected optimal $p$ larger than one indicates that emphasizing spectral peaks improves model fit with respect to standard Whittle estimation. Hence, the time series is likely characterized by strong periodicities or oscillations (including possibly long-range dependence if the peak is at low frequency), as the optimal $p$ lets estimation focus on capturing high periodogram ordinates. On the other hand,

optimal $p < 1$ downweights spectral peaks, suggesting they are spurious and preventing them from contaminating and distorting model fit. When salience enters the series via its surprise and prominence features, some characteristics of the series are upweighted and others are discounted. The parameter is selected based on a goodness-of-fit weighted Whittle-type criterion, and inference is conducted by means of bootstrap procedures.

The contribution of the paper is threefold. First, it provides a statistical characterization of salience for temporally dependent data, thereby extending the salience framework from cross-sectional settings to time series. Second, it develops an operational estimation strategy that can be applied to observed historical data. Furthermore, it shows how the optimally selected salience parameter can be used to study expectation formation in macroeconomics.

The empirical application focuses on U.S. inflation and household inflation expectations. This setting is particularly well suited to the analysis. Inflation is one of the main macroeconomic variables through which households observe the economy, and inflation expectations are central to models of consumption, wage setting, and monetary policy transmission. At the same time, recent inflation dynamics–especially the post-2021 inflation surge–raise a natural question as to whether expectations incorporate macroeconomic information continuously or whether they respond disproportionately to particularly salient developments.

To address this issue, we select the optimal salience parameter for three related series: realized inflation, aggregate inflation expectations, and forecast errors. This comparison allows us to distinguish between the structure of the underlying macroeconomic signal, the part internalized by beliefs, and the residual component not incorporated into expectations. The empirical results show that all three series display values of $p$ above unity, indicating strong spectral structure. Realized inflation exhibits the highest value, aggregate expectations also display substantial structure, and forecast errors retain residual dynamics. Taken together, the findings suggest that the dynamic structure of inflation is only partially incorporated into expectations, consistent with models in which agents process macroeconomic information selectively rather than under continuous full-information

updating.

The paper is related to two strands of literature. The first is the literature on salience and diagnostic expectations, which emphasizes that individuals overweight prominent features of the environment when forming beliefs. The second is the literature on inflation expectations, which studies the extent to which household beliefs track realized inflation and which cognitive or informational frictions shape forecast behavior. Our contribution differs from both strands in that it does not focus on cross-sectional heterogeneity in respondents, but on the temporal structure of the aggregate series itself.

The remainder of the paper is organized as follows. Section 2 describes salience theory and provides a literature overview which helps clarify the positioning of this paper. Section 3 introduces the notion of salience in a time series setting and discusses its relation to surprise and prominence. Section 4 develops the spectral framework and the estimation method. Section 5 presents illustrative examples, while Section 6 contains simulation evidence. Section 7 applies the methodology to U.S. inflation, inflation expectations, and forecast errors. Section 8 concludes.

# 2 Literature

Salience theory, introduced by Bordalo, Gennaioli, and Shleifer, provides a psychologically grounded framework in which attention is endogenously directed toward features of the environment that stand out relative to a reference or comparison object. In their foundational work, Bordalo et al. (2012) develop a model of choice under risk in which decision makers overweight outcomes that are most salient—i.e., those that differ most from alternative outcomes—leading to context-dependent probability distortions and systematic deviations from expected utility theory. This approach is extended to financial markets in Bordalo et al. (2013), where asset prices reflect the overweighting of salient payoffs, generating mispricing driven by attention to extreme outcomes. More generally, salience theory implies that individuals do not process all available information symmetrically, but instead focus on features that are particularly prominent relative to a benchmark.

Subsequent work integrates salience with a theory of memory and reference points. (Bordalo et al., 2020) develop a model in which attention is shaped by surprise relative to a memory-based norm, itself constructed through associative recall of past experiences. In this framework, economic behavior depends on both the formation of norms and the degree to which current outcomes deviate from them. Relatedly, Bordalo et al. (2019) show how memory-based reference prices affect economic decisions such as rental choice, generating systematic deviations from standard theory that gradually disappear with experience.

Taken together, this literature highlights two central mechanisms: (i) the formation of reference points or norms through past experience, and (ii) the overweighting of outcomes that are particularly different from those norms. These mechanisms provide a unified explanation for a wide range of behavioral phenomena, including context effects, instability of preferences, and gradual adaptation to new environments.

The salience-based view of attention is closely related to the broader literature on bounded rationality and limited information processing. Rational inattention models (e.g., Sims, 2003; Miao et al., 2022) formalize optimal allocation of attention across signals, while alternative approaches emphasize heuristic or state-dependent reasoning. In particular, Ilut and Valchev (2023) model agents as imperfect problem solvers whose reasoning becomes more active in unusual or complex states, generating endogenous variation in attention.

A complementary empirical literature studies how attention affects macroeconomic expectations. Briand et al. (2025) document that measures of attention to inflation—such as Google searches and media coverage—help explain both inflation dynamics and expectations, especially during high-inflation periods. These findings suggest that attention is a key channel through which macroeconomic developments are transmitted into beliefs.

A smaller but growing literature attempts to operationalize salience empirically outside laboratory settings. Parsley and Popper (2024), for example, construct a measure of climate-change salience based on unexpected spikes in public attention and show that it is priced in international equity markets. This approach highlights the importance of

distinguishing between persistent attention and unexpected, salient shocks.

Balcombe et al. (2021) provide an econometric assessment of salience theory, emphasizing that its empirical implementation requires careful measurement of which features of the environment are effectively "salient" to agents. These contributions underscore the need for reduced-form measures of salience that can be directly estimated from data.

Our analysis contributes to the literature by proposing a time-series-based, frequency-domain measure of salience. Rather than identifying salient events ex ante or relying on external proxies for attention, we infer the degree of salience directly from the stochastic structure of economic time series. The key parameter $p$ captures the extent to which variability is concentrated in a subset of spectral components, providing a reduced-form measure of how strongly the dynamics of a series depart from a diffuse, noise-like benchmark.

Applying this methodology to realized inflation, inflation expectations, and forecast errors, we provide evidence on how the dynamic structure of macroeconomic variables is transmitted into beliefs. In doing so, the paper bridges behavioral theories of salience and attention with empirical macroeconomic analysis, offering a novel perspective on expectation formation and information processing.

# 3 Time series salience

In this Section, we provide a characterization of salience for temporally dependent data. This is substantially different from the setting considered by Bordalo et al. (2022) for cross-sectional independent data and this different characterization translates into distinct effects of salience on individuals' behavior.

First, note that salience characteristics of given economic variables described in Bordalo et al. (2022) are shown to affect consumers' economic choice and choice under risk. In the time series setting that we consider, the decision problem corresponds to a forecasting problem, and the effects of salience show up in the individuals' selection of the next period's forecast for the temporally dependent variable of interest.

We next translate the three features of salience outlined by Bordalo et al. (2022) into time series features and use the description we provide to analyze how such features affect the forecasts by shaping individuals' attention.

As already mentioned, the three factors shaping salience are *i*) contrast with surroundings, *ii*) surprise, or contrast with past, and *iii*) prominence. It should be noted that here we consider univariate time series, and focus only on factors *ii*) and *iii*) as a contrast with surroundings corresponds to contrast with past, captured by surprise.

## 3.1 Surprise as a Dimension of Salience in Time Series

In the context of time series analysis, *surprise* represents a key dimension of salience and refers to the extent to which observed realizations deviate unexpectedly from recent past behavior or from the underlying data generating mechanism. Unlike simple variability, surprise captures *localized, episodic, or abrupt changes* in the temporal or spectral structure of the process that attract the observer's attention.

From a time–domain perspective, surprise arises when the series exhibits sudden changes in its dynamics, such as regime shifts, temporary oscillatory behavior, or abrupt variations in amplitude. These features generate strong contrasts with respect to past observations, making them particularly salient to an observer who forms expectations based on recent data. Importantly, surprise does not require persistent dependence or long-range memory; rather, it is driven by *unexpected departures from regularity*, even in globally short–memory processes.

In the frequency domain, surprise manifests itself as *localized peaks in the sample spectrum*. These peaks correspond to frequency components that are temporarily or irregularly activated and are not uniformly present across time. As a result, the sample spectrum displays marked discrepancies with respect to a smooth or reference spectral model, reflecting the presence of salient features concentrated at specific frequencies rather than a global redistribution of power.

### 3.1.1 Illustrative examples: the surprise component of salience

The surprise component of salience refers to the presence of observations or patterns that deviate markedly from recent or habitual behavior, in a way that is immediately noticeable to an observer. In the context of historical data, surprise does not necessarily require nonstationarity or permanent structural breaks. Rather, it may arise from transient but pronounced features that stand out against the background variability of the series.

In what follows, we provide two illustrative examples. Their purpose is descriptive and intuitive: they are meant to clarify how surprise manifests itself in observed data.

**Example 1: Surprise in a single historical series.** We consider a univariate time series whose overall behavior is covariance–stationary and short–memory, but which occasionally displays episodes of unusually strong and coherent oscillations. The data are generated from a stationary process obtained as the sum of a weak background disturbance and a cyclical component that becomes intermittently dominant over limited time intervals.

More precisely, the cyclical component $s_t$ follows a stationary AR(2) model with complex roots,

$$s_t = \phi_1 s_{t-1} + \phi_2 s_{t-2} + \eta_t, \qquad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2),$$

with $\phi_1 = 2r\cos(\omega_0)$ and $\phi_2 = -r^2$, where $0 < r < 1$ controls damping and $\omega_0$ fixes the cycle frequency. To obtain *localized* surprise episodes without introducing long-run dependence, the observed series is constructed as

$$y_t = g_t\, s_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2),$$

where $g_t \in \{0, 1\}$ is a block–wise ON/OFF indicator that activates the cyclical component only over short time intervals. This mechanism produces bursts of regular oscillations that stand out against the surrounding fluctuations, while keeping the overall process stationary.

Surprise, in this example, does not arise from permanent changes in persistence or

long memory, but from the temporary emergence of a distinctive dynamic pattern that contrasts sharply with recent observations. An observer following the series sequentially would perceive these bursts as unexpected deviations from the local regularity characterising most of the sample.

Figure 1 reports in the top panel the simulated series in levels, highlighting the episodic oscillatory behavior. The bottom panel shows the corresponding first differences, where surprise episodes translate into sequences of unusually large and regular changes. Figure 2 reports the raw periodogram, which displays a pronounced peak around $\omega_0$ associated with the activated cycle.

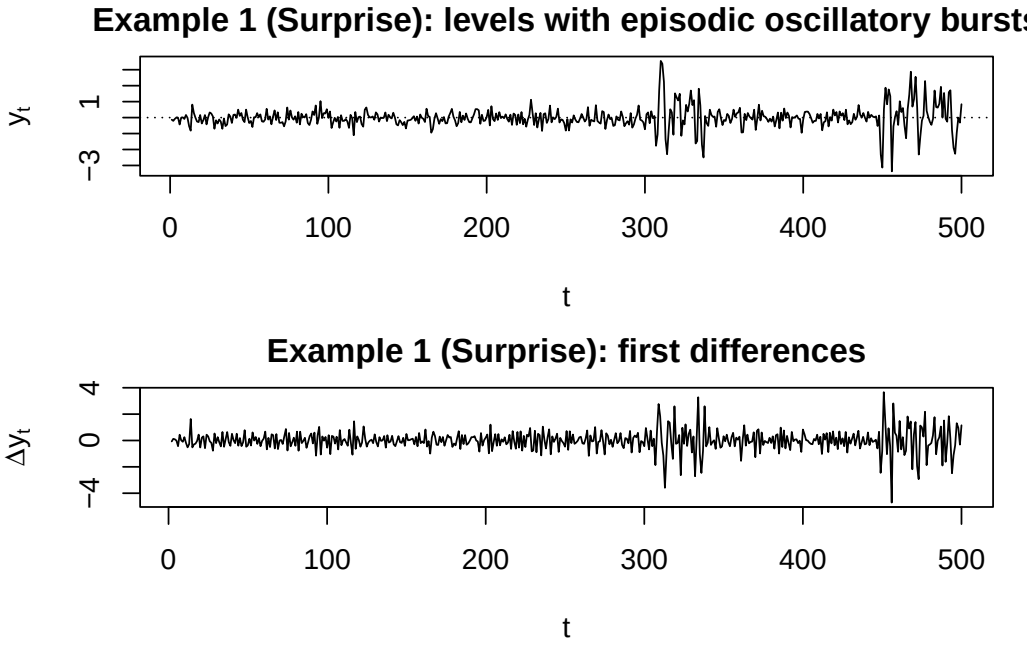



Figure 1: Example 1. Top: simulated series in levels $y_t$. Bottom: first differences $\Delta y_t$.

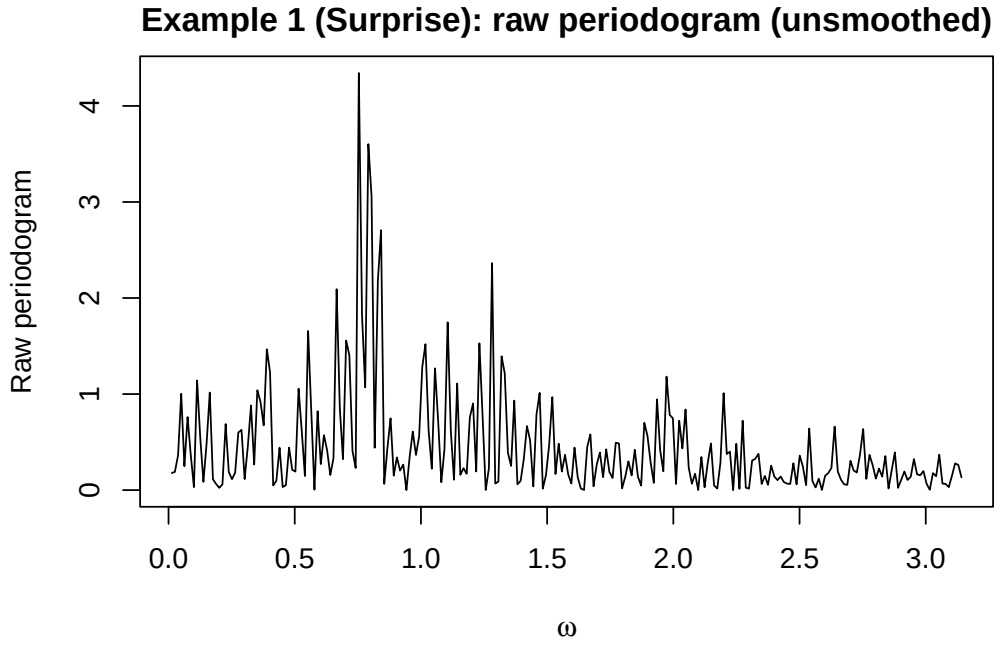



Figure 2: Example 1: raw periodogram showing a peak around the cycle frequency $\omega_0$.

**Example 1: surprise in a multivariate system and comparative statics** Surprise can also be understood in a multivariate or system-based setting, even when attention is ultimately focused on individual series. In many economic contexts, observed variables are jointly determined and respond to common shocks or structural changes. From this viewpoint, surprise corresponds to a sudden shift in the configuration of the system, moving it from one regime or equilibrium to another.

To illustrate this idea, consider a stationary VAR(1) system collecting three economic variables (for example, output, inflation, and an interest rate),

$$y_t = c + Ay_{t-1} + u_t, \qquad u_t \sim \text{i.i.d.}(0, \Sigma),$$

where the eigenvalues of $A$ lie inside the unit circle, so that the process is covariance-stationary within each regime.

A *comparative statics* experiment is constructed by generating two trajectories under the *same sequence of innovations* $\{u_t\}$. The first trajectory represents a counterfactual baseline in which parameters remain constant over the full sample. The second trajectory corresponds to a policy-change scenario in which, at a known date $T_0$, the system switches

from $(A_{\text{pre}}, c_{\text{pre}})$ to $(A_{\text{post}}, c_{\text{post}})$:

$$y_t = \begin{cases} c_{\text{pre}} + A_{\text{pre}} y_{t-1} + u_t, & t < T_0, \\ c_{\text{post}} + A_{\text{post}} y_{t-1} + u_t, & t \geq T_0. \end{cases}$$

Because both scenarios are driven by identical innovations, any divergence observed after $T_0$ is entirely attributable to the structural change in the system rather than to random fluctuations. The comparison therefore isolates the effect of the change in the data generating mechanism.

Figure 3 plots each variable under the policy-change scenario (solid line) and under the counterfactual baseline (dashed line), together with a vertical line marking $T_0$. Before the transition, the two paths coincide closely; after $T_0$, the policy-change trajectory departs visibly from the counterfactual. This before–after contrast makes the change immediately salient to an observer.

In this example, surprise is not associated with isolated shocks or outliers, but with an abrupt and unanticipated change in the system's structure that alters the qualitative behavior of the variables. Importantly, the series remain stationary before and after the change. Surprise therefore arises from a shift in expectations induced by the structural transition, rather than from increased persistence or long memory, which are instead related to the prominence dimension of salience.

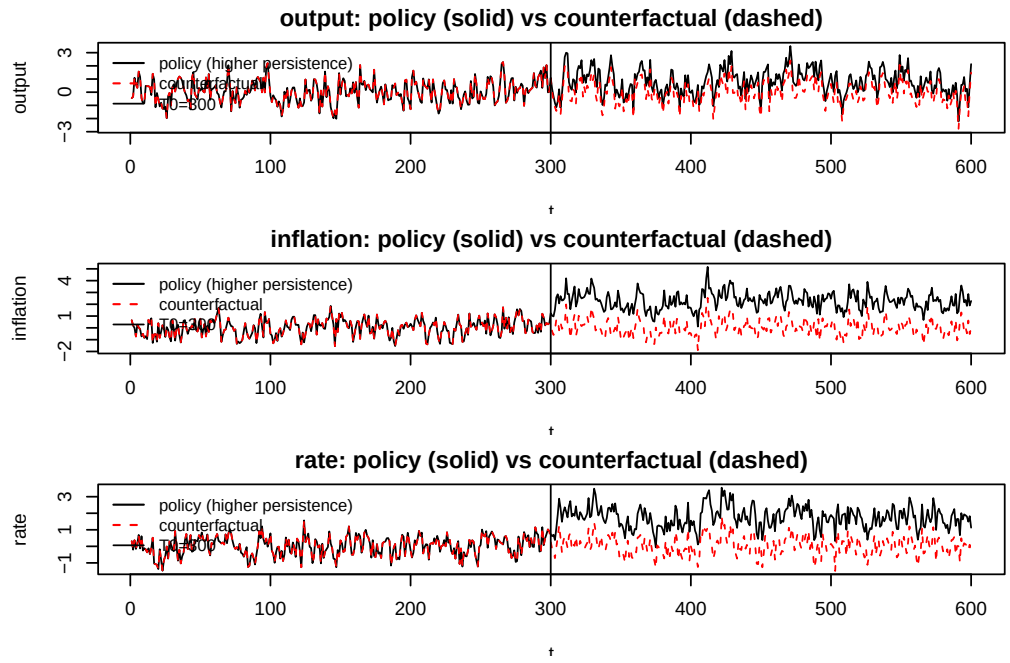



Figure 3: Example 1: comparative statics in a VAR system: policy-change scenario (solid line) versus counterfactual baseline (dashed line). The vertical line indicates the transition date $T_0$.

## 3.2 Prominence as a Dimension of Salience in Time Series

In our statistical translation, *prominence* corresponds to persistence in the data and is operationalised through long memory behavior. Prominence reflects the degree to which patterns in a time series are repeatedly observed over time, making them cognitively more available to the observer. Increasing the number of opportunities to observe a given change—either positive or negative—in the series enhances its salience through repetition rather than through unexpected deviations.

An individual observing subsequent increments (or decrements) in the levels of a time series will tend to predict that the series is going to increase (decrease) further in the next period. Successive changes of the same sign produce local trends, reinforcing expectations about future behavior. More generally, when a time series is characterized by long memory, its realizations at a given time point exert a persistent influence on future values, so that past observations remain informative over long horizons.

From a visual perspective, prominent time series display extended intervals during which observations remain at relatively high or low levels. These local trends emerge even

in the absence of abrupt changes and contribute to salience by continuously reinforcing the same directional signal. Unlike surprise, which is driven by sudden or episodic deviations, prominence is cumulative in nature and arises from sustained dependence over time.

In the frequency domain, prominence manifests itself as an accumulation of spectral power at low frequencies. Long memory processes are characterized by a spectral density that diverges near the zero frequency, producing a pole that reflects the dominance of slowly evolving components. As a consequence, the sample spectrum exhibits systematically large ordinates in the low–frequency region.

Within the proposed framework, such persistent spectral features affect estimation through the power parameter $p$. When prominence is present, emphasizing low–frequency discrepancies between the sample and theoretical spectrum improves model fit. This leads the weighted Whittle deviance criterion to select values $p > 1$, which assign greater importance to dominant and repeatedly observed components of the spectrum. Hence, large selected values of $p$ indicate that persistence–driven salience plays a central role in shaping the second–order structure of the data.

Taken together, prominence captures a form of salience rooted in memory and repetition rather than in novelty. While surprise highlights deviations that stand out because they are unexpected, prominence highlights features that stand out because they are repeatedly encountered. The power parameter $p$ provides a unified way to detect both mechanisms, with prominence primarily influencing the weighting of low–frequency components.

### 3.2.1 Illustrative examples: the prominence component of salience

The prominence component of salience is related to *persistence* and refers to the extent to which observed patterns remain available to the observer because they last over time. In historical data, prominence manifests itself through long stretches of similar behavior, local trends, and slow mean reversion. Unlike surprise, prominence is not driven by abrupt deviations, but by the *continuity* of the dynamics, which makes recent configurations more cognitively accessible.

**Example 2: prominence in a single historical series** We first consider a single stationary historical series exhibiting high prominence. The data are generated from a stationary long memory specification, such as an ARFIMA$(0, d, 0)$ model with $d \in (0, 0.5)$. For a fixed (high) value of $d$, realizations display persistent deviations from the mean and long runs of similar behavior, which are visually perceived as local trends. Figure 4 shows the series in levels. Figure 5 reports the smoothed estimated spectral density, where prominence translates into a strong concentration of power at low frequencies, reflecting slow-moving components.

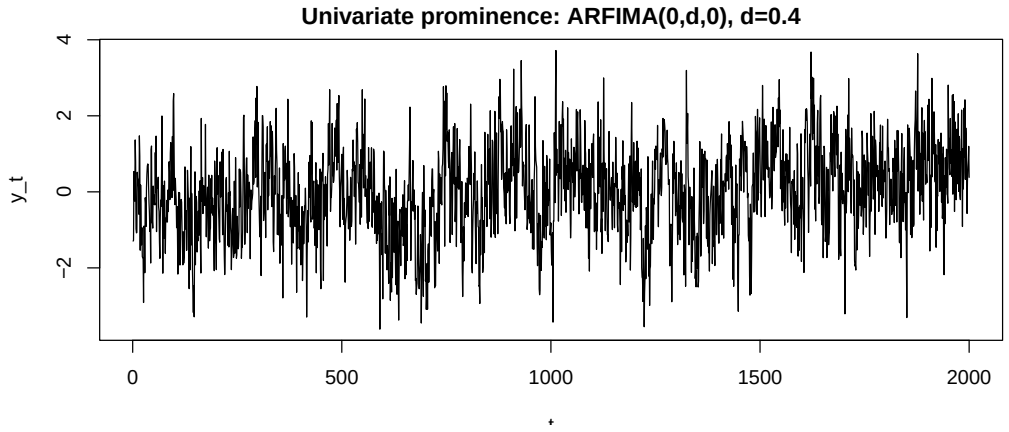



Figure 4: Prominence in a univariate stationary series: persistent behavior in levels.

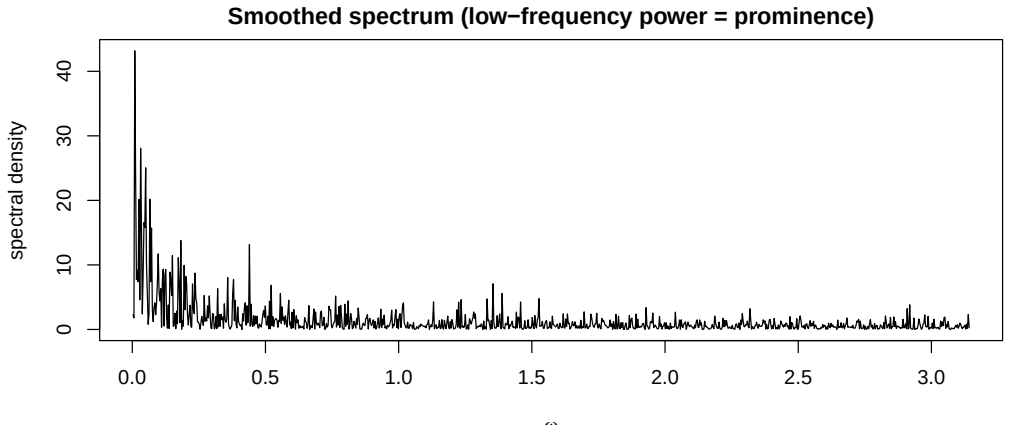



Figure 5: Smoothed spectrum: prominence corresponds to higher power at low frequencies.

### 3.2.2 Example 2: Prominence in a multivariate system and comparative statics

Prominence can also be illustrated in a multivariate setting, where individual time series are embedded in a system of jointly determined variables. In many economic environments, variables such as output, inflation, and interest rates evolve according to common dynamics and respond to shared shocks. In this context, prominence is associated with the *persistence of these dynamics*, which makes observed patterns more temporally extended and therefore more cognitively available to an observer.

To illustrate this idea, consider a stationary VAR(1) system collecting three economic variables,

$$y_t = c + Ay_{t-1} + u_t, \qquad u_t \sim \text{i.i.d.}(0, \Sigma),$$

where the eigenvalues of $A$ lie inside the unit circle. Within each regime, the system is covariance–stationary and exhibits stable second–order properties.

The prominence feature is highlighted through a *comparative statics* experiment. Two scenarios are generated using the *same sequence of innovations* $\{u_t\}$. In the counterfactual baseline, the system parameters remain constant over the entire sample. In the alternative scenario, at a known date $T_0$, the system undergoes a structural change in its propagation mechanism: the autoregressive matrix switches from $A_{\text{pre}}$ to a more persistent matrix $A_{\text{post}}$, while the intercept term remains unchanged. Formally,

$$y_t = \begin{cases} c + A_{\text{pre}} y_{t-1} + u_t, & t < T_0, \\ c + A_{\text{post}} y_{t-1} + u_t, & t \geq T_0, \end{cases}$$

with both matrices satisfying stationarity conditions.

Because the two scenarios share the same innovations, any divergence observed after $T_0$ is entirely attributable to the change in the system dynamics. Before the transition, the two trajectories are nearly indistinguishable. After $T_0$, the series generated under the more persistent regime display longer–lasting responses to shocks and smoother, more inertial movements over time.

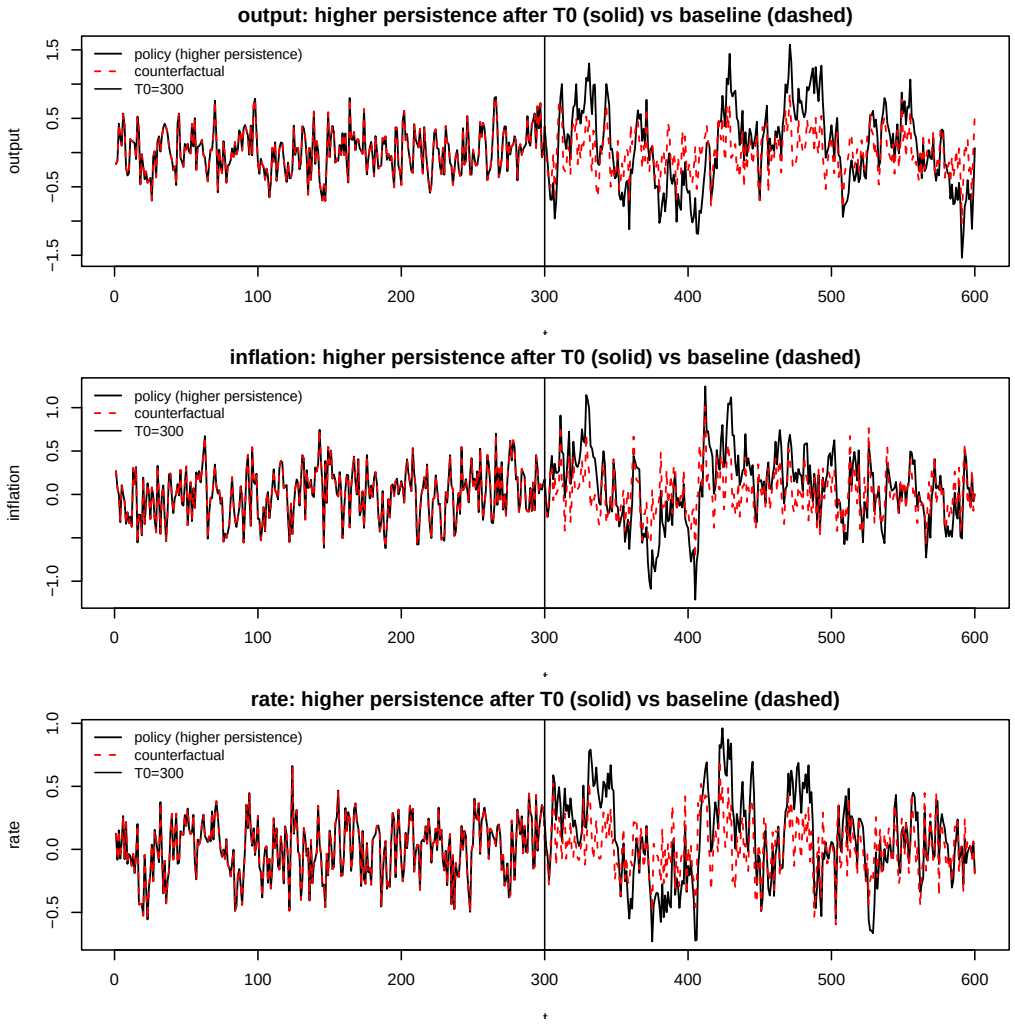



Figure 6: Example 2: Prominence in a VAR(1) system via comparative statics. Solid line: policy-change scenario (at $T_0$). Dashed line: counterfactual baseline with constant parameters. The vertical line marks the transition date $T_0$.

This example illustrates prominence as a salience feature arising from persistence rather than from abrupt deviations. Unlike surprise, which is driven by localized and unexpected changes, prominence reflects the tendency of the system to reinforce and propagate past realizations, thereby increasing the visibility and cognitive availability of patterns over time.

# 4 Methods

We propose statistical tools to detect time series properties corresponding to the above mentioned salience characteristics. If salience is present in the data, this causes certain

aspects of the time series to be highlighted and others to be discounted by the observer, based on how strongly they attract her attention.

## 4.1 Functionals of the spectrum and the power process

The temporal dependence structure of stationary stochastic processes can be either analyzed in the time domain, via the autocovariance function, or in the frequency domain, based on the spectral density function. Frequency domain analysis or spectral analysis provides an alternative way of viewing the process, which may be more illuminating for some applications. It is based on the decomposition of the process into the sum of sinusoidal components with uncorrelated random coefficients. The second-order properties of a stationary stochastic process $\{X_t\}_{t\in T}$ are described by its autocovariance function (ACVF), $\gamma_k(\theta) = \mathrm{E}[X_t X_{t-k}]$, depending on the $L \times 1$ vector of unknown parameters $\theta = (\theta_1, \ldots, \theta_L)^\top \in \Theta \subseteq \mathbb{R}^L$. The Fourier transform of the ACVF allows to translate the autocovariance structure in the frequency domain. The spectral density function of $\{X_t\}$ is the function $f(\cdot;\theta)$ defined by the Fourier transform of the ACVF

$$f(\lambda;\theta) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k(\theta) \exp(-i\lambda k), \quad -\pi \leq \lambda \leq \pi. \tag{1}$$

In turn, the ACVF can be recovered by the inverse Fourier transform,

$$\gamma_k(\theta) = \int_{-\pi}^{\pi} f(\lambda;\theta) \exp(i\lambda k)\, d\lambda. \tag{2}$$

The spectral density function provides the contribution of each frequency component to the total variability of the process. Taking powers of it allows to emphasize certain characteristics of the process. The variance profile (Luati et al., 2012) and the generalised autocovariance function (Proietti and Luati, 2015) are defined as functionals of a power transformation of the spectral density of a stationary stochastic process. Specifically, let $\{X_t\}_{t\in T}$ be a stationary zero-mean stochastic process with Wold representation $X_t = \Psi(B)\varepsilon_t = (1 + \Psi_1 B + \Psi_2 B^2 + \ldots)\varepsilon_t, \quad \sum_j \psi_j^2 < \infty, \varepsilon_t \sim i.i.d.(0, \sigma^2)$, where $B^k \varepsilon_t = \varepsilon_{t-k}$. Its spectral density function is $f(\omega) = \frac{1}{2\pi}[\Psi(e^{-i\omega t})\Psi(e^{i\omega t})]$.

By tuning the power parameter $p$ that governs the transformation of the spectrum, the variance profile $v_p$ provides useful quantities characterising the second-order properties of the process. $v_p$ can be interpreted in terms of a power transformation $u_{pt}$ of the original process $x_t$. The power process is defined as

$$u_{pt} = \begin{cases} \Psi(B)^p \varepsilon_t \quad \text{for} \quad p \geq 0 \\ \Psi(B^{-1})^p \varepsilon_t \quad \text{for} \quad p < 0 \end{cases} \tag{3}$$

Note that for $p = 1$ the variance profile coincides with the unconditional variance of the process, and $u_{1t} = x_t$, while for $p \to 0, u_{0t} = \varepsilon_t$ and $v_p$ converges to $\sigma^2$, the one-step-ahead prediction error variance.

The power process (3) is also linked to the generalised autocovariance (GACV) function, defined in Proietti and Luati (2015). The underlying idea is that taking powers of the spectral density is useful to detect certain features of the time series. For instance, $p > 1$ emphasizes peaks in the spectral density, while negative values of $p$ highlight troughs. The GACV can be interpreted as the ACVF of the power process $u_{pt}$ in (3). For $p = 1$ the GACV coincides with the traditional ACVF, and $u_{1t} = x_t$.

The definition of the power process is key to our purposes, as discussed next.

The surprise and prominence factors shaping salience, show up in time series data as described in the previous Section. They consist of sudden and marked changes that lead to unexpected departures from background regularity, and in the repetition of changes of the same sign, visually producing the local trends that characterize strong persistence in the data. Such behavioral patterns produced by surprise and prominence, affect the shape of the spectrum of the underlying process in distinct ways.

Surprise translates into *irregular and concentrated peaks* of the sample spectrum at specific frequencies. Such peaks may arise from temporary oscillatory behavior, sudden changes in amplitude, or abrupt regime shifts that are not persistent over time. Impor-

tantly, surprise does not imply a systematic redistribution of power across frequencies. Instead, it generates discrepancies between the sample spectrum and a smooth reference spectrum that are confined to selected frequency bands. As a result, the spectral density remains globally short–memory, but exhibits salient local features that stand out relative to the surrounding frequencies.

Prominence, by contrast, is linked to *persistent and recurrent patterns* that remain visible over long horizons. In the frequency domain, prominence manifests itself as a *structural concentration of power at low frequencies*. This behavior is typical of long memory or highly persistent processes, where observations at distant time points remain strongly correlated. Spectrally, this persistence is reflected in a slowly decaying spectral density and, in the case of fractional integration, in the presence of a pole at the zero frequency. Unlike surprise, prominence affects the spectrum in a global and systematic manner, shaping its overall form rather than generating isolated irregularities.

We consider a general class of spectral models defined by Proietti and Luati (2015), based on the GACV. By tuning a parameter $p \in \mathbb{R}$ which defines the power transform of the spectrum of a time series, this model encompasses both the AR and MA models.

$$2\pi f_p(\omega) = \left[ \frac{\sigma_p^2}{\phi_p(e^{-i\omega})\phi_p(e^{i\omega})} \right]^{\frac{1}{p}}, \tag{4}$$

where, for $p > 0$ ($p < 0$) $\phi_p(z) = 1 - \phi_{p1}z - \phi_{p2}z^2 - \cdots - \phi_{pK}z^K$ is the characteristic polynomial of the AR (MA) approximation of $u_{pt}$. For a given $p$, the model can be estimated as suggested by the authors, solving a Yule-Walker system of equations in the GACV nonparametric estimator. Then, the optimal power transform $p$ can be selected over a specified grid by optimizing a measure of goodness-of-fit. Distinct values of $p$ emphasize certain characteristics of the process, and can be useful to uncover the presence of the above mentioned salient features in the time series of interest.

Within the proposed framework, these two dimensions of salience interact differently with

estimation. Surprise induces sharp, frequency–specific mismatches between the empirical and theoretical spectra, while prominence produces broad and persistent deviations concentrated at low frequencies. The power parameter $p$ governs how strongly such discrepancies are emphasized: values $p > 1$ assign greater weight to frequencies where the mismatch is largest, enhancing sensitivity to both localized spectral peaks associated with surprise and low–frequency accumulations associated with prominence. In this sense, the spectral density provides a natural domain in which the distinct manifestations of salience can be identified and quantitatively assessed.

Proietti and Luati (2015) suggest to select $p$ by minimizing a deviance measure equal to minus twice the Whittle likelihood. Here we find it useful to consider Whittle likelihood estimation, as it allows interpretation of the power parameter in terms of the salience of the series.

## 4.2 Whittle likelihood estimation

If one considers estimation of the general spectral model (4) by solving the Yule-Walker equations in the nonparametric GACV estimator, this is equivalent to Whittle estimation of an AR($K$) model for $u_{pt}$ (Proietti and Luati, 2015).
Given a sample $X_1, \ldots, X_n$ from the process $\{X_t\}$, the Whittle likelihood function (Whittle, 1953) is defined as

$$\ell_n(\theta) = -\sum_{\omega \in \Omega_n} \left\{ \log f(\omega;\theta) + \frac{I_n(\omega)}{f(\omega;\theta)} \right\}, \tag{5}$$

where $\Omega_n$ denotes the set of Fourier frequencies

$$\Omega_n = (\omega_1, \ldots, \omega_n) = \frac{2\pi}{n}(1, \ldots, \lfloor (n-1)/2 \rfloor). \tag{6}$$

We focus on fitting a spectral model $f_u(\omega)$ to the process $u_{pt}$, where $2\pi f_u(\omega) = [2\pi f_p(\omega)]^p$. Then, given $n$ observations from $\{X_t\}$, the Whittle likelihood we want to optimize is

$$\begin{aligned}\ell_u(\theta, p) &= -\sum_{\omega\in\Omega_n}\left\{\log[f_u(\omega;\theta)] + \frac{[2\pi I_n(\omega)]^p}{f_u(\omega;\theta)}\right\} = \\ &= -\sum_{\omega\in\Omega_n}\left\{\log\{[2\pi f_p(\omega)]^p\} + \frac{[I_n(\omega)]^p}{[2\pi f_p(\omega)]^p}\right\} \qquad (7)\end{aligned}$$

As previously stated, estimation by the generalised Yule-Walker equations, with subsequent optimal selection of $p$ is equivalent to Whittle estimation by means of equation (7). This leads to the interpretation of estimation as a weighted optimization problem, where the second-order characteristics of the observed series are given different weights, depending on the parameter $p$.

To see this, consider the score function corresponding to the traditional Whittle likelihood $\ell_n$ and that corresponding to $\ell_u$, both fitting the spectral model $f_p(\omega;\theta)$ to $x_t$ and $u_{pt}$ respectively:

$$\nabla_\theta \ell_n(\theta) = S_n = \sum_\omega s(\omega;\theta), \qquad (8)$$

$$s(\omega;\theta) = \nabla_\theta f_p(\omega;\theta)\left[\frac{I(\omega)}{f_p(\omega)^2} - \frac{1}{f_p(\omega)}\right] \qquad (9)$$

$$\nabla_\theta \ell_u(\theta, p) = S_u = p\sum_\omega s_u(\omega;\theta,p), \qquad (10)$$

$$s_u(\omega;\theta,p) = \nabla_\theta f_p(\omega;\theta)\left[\frac{I(\omega)^p}{f_p(\omega)^{p+1}} - \frac{1}{f_p(\omega)}\right] \qquad (11)$$

Note that $S_u$ can be expressed as a weighted sum of the individual contributions to the standard Whittle likelihood score, with weights depending on the sample and theoretical spectra and the power transformation $p$ :

$$S_u = p \sum_{\omega} w_p(\omega;\theta) s(\omega;\theta), \tag{12}$$

$$w_p(\omega;\theta) = \frac{\left( \frac{I(\omega)^p}{[f_p(\omega;\theta)]^{p+1}} - \frac{1}{f_p(\omega;\theta)} \right)}{\left( \frac{I(\omega)}{f_p(\omega;\theta)^2} - \frac{1}{f_p(\omega;\theta)} \right)} \tag{13}$$

The weights control the score contributions at each frequency to the total score, depending on the model fit to the observed spectrum.

Hence, in a similar fashion as Zanetti Chini (2023) in the context of measuring judgement, $p$ provides information about the salience of the series. Specifically, if $p > 1, w_p(\omega;\theta)$ emphasizes mismatches such that estimation is more sensitive to periodogram peaks.

The optimal value of $p$ over the chosen grid of values is the one that maximizes the fit of the spectral model to the sample spectrum not uniformly, but frequency-wise, depending on $w_p(\omega;\theta)$. Selected values $\widehat{p} > 1$ indicate that emphasizing spectral peaks improves model fit with respect to standard Whittle estimation. Hence, the time series is likely characterized by strong periodicities or oscillations (including also strong persistence and possibly long-range dependence if the peak is at low frequency), as the optimal $p$ lets estimation focus on capturing high periodogram ordinates. On the other hand, estimated $\widehat{p} < 1$ downweights spectral peaks, suggesting they are spurious and preventing them from contaminating and distorting model fit. This leads to more robust estimation.

When salience enters the series via its surprise and prominence features, some characteristics of the series are upweighted and others are discounted.

## 4.3 Bootstrap inference for the salience parameter

Because the salience parameter $p$ is selected over a discrete grid rather than obtained from a continuous optimization problem, its sampling variability cannot be assessed using standard asymptotic results for smooth estimators. Instead, we treat $\widehat{p}$ as an *argmin*

*estimator* obtained from the minimization of a sample criterion,

$$\widehat{p} = \arg\min_{p \in \mathcal{P}} \mathrm{dev}_n(p),$$

where $\mathrm{dev}_n(p)$ denotes the empirical deviance constructed from the smoothed spectral estimate. Inference is therefore based on a bootstrap procedure that replicates the entire criterion function and recomputes the minimizer in each resample. This approach follows the general bootstrap principles for argmin-type estimators discussed in Van Der Vaart and Wellner (1996).

Operationally, the procedure is as follows. First, the periodogram of the observed series is computed and a pilot spectral density estimate is obtained from the selected model. The periodogram ordinates are then standardized by this pilot spectrum and resampled with replacement in the frequency domain, producing bootstrap pseudo-periodograms that preserve the second-order dependence structure of the data. For each bootstrap replicate, a pseudo-series is constructed solely for the purpose of obtaining a smoothed spectral estimate, and the deviance function $\mathrm{dev}_n^*(p)$ is recomputed over the full grid of candidate values. The bootstrap estimate $p^*$ is then defined as

$$p^* = \arg\min_{p \in \mathcal{P}} \mathrm{dev}_n^*(p).$$

Repeating this procedure yields a bootstrap sample $\{p_b^*\}_{b=1}^B$ that approximates the sampling distribution of $\widehat{p}$. From this distribution we compute bootstrap summaries such as the mean, bias, and percentile confidence intervals for $p$.

This resampling strategy provides an empirical measure of uncertainty for the salience parameter and allows one to evaluate the robustness of the salience classification (for example whether $p > 1$) across resamples. By operating in the frequency domain and resampling standardized periodogram ordinates, the procedure preserves the dependence structure of the series while remaining compatible with the spectral framework underlying the proposed salience estimator.

## 4.4 Assumptions

In this work we consider a stationary zero-mean stochastic process with Wold representation $X_t = \Psi(B)\varepsilon_t = (1+\Psi_1 B+\Psi_2 B^2+\dots)\varepsilon_t, \quad \varepsilon_t \sim I.I.D.(0,\sigma^2)$, where $B^k\varepsilon_t = \varepsilon_{t-k}$. Its spectral density function is $f(\omega;\theta) = \frac{\sigma^2}{2\pi}|\Psi(e^{-i\omega})|^2$, where $\theta \in \Theta \subseteq \mathbb{R}^L$ is the vector including all model parameters to be estimated. In this setting, $p$ is not a model parameter but a tuning parameter of the spectral transformation. It is selected separately, to optimize a measure of goodness-of-fit. We impose the following assumptions:

C1 : $\sum_{j=0}^{\infty} j^{\delta}|\Psi_j| < \infty, \delta > \frac{3}{4}$;

C2 : $\mathrm{E}[|\epsilon_t|^s] = \mu_s < \infty$, where $s > max\{9, 4(p+m), 8(p-1)+2m, 2(p-1)+4m\}$;

C3 : $\sup_{|\theta|\geq\theta_0} |\mathrm{E}[e^{i\theta\epsilon_t}]| = \delta(\theta_0) < 1$;

C4 : $\int_{-\pi}^{\pi} |\mathrm{E}[e^{i\theta\epsilon_t}]|^r \, d\theta < \infty$ for some $r \geq 1$;

C5 : $m$ is a fixed positive integer satisfying $m + 4(p-1) > 0$.

The above are technical assumptions from Proietti and Luati (2015) to ensure applicability of the method and good properties of the estimator proposed by the authors.

The theoretical results hold under the above stated assumptions, for covariance–stationary processes. In empirical applications, however, surprise–driven salience may arise precisely through nonstationary behavior, such as level shifts, changing volatility, or temporary episodes of markedly different dynamics. In this case, the proposed methodology can still be implemented after a transformation that restores (approximate) stationarity and isolates the salient component of interest. In practice, this can be achieved by (i) removing deterministic components (intercept, linear trend, seasonality), (ii) applying differencing or log–differencing to eliminate stochastic trends, and/or (iii) working with deviations from a smooth benchmark (e.g. residuals from an ARMA/VAR model or from a low–frequency filter) so that the remaining series captures the unexpected, localized departures from regular behavior. The power–Whittle selection of $p$ is then applied to

the transformed series, which is designed to retain the high–contrast features that define salience while discarding nonstationary drift. Importantly, this procedure does not "remove" salience: rather, it separates salience from slow–moving components that are unrelated to attention–grabbing events. If surprise manifests as abrupt changes or episodic oscillations, these features typically remain visible after stationarising transformations and continue to generate localized discrepancies in the sample spectrum. Consequently, the selected $p$ can still exceed unity, indicating that the transformed likelihood assigns relatively higher weight to those frequency regions where the salient deviations concentrate. In this sense, a value $p > 1$ computed on a stationarised representation of the data can be interpreted as evidence that salient features are present in the underlying observations, even when the original series is nonstationary.

# 5 Examples

Let us consider simple analytical examples to gain a deeper insight into the role of the power parameter $p$ in model estimation and its interpretation relative to the salience of the series.

## 5.1 General Example

We start with a generic example, leaving the spectral model $f(\omega, \theta)$ unspecified. In equation (7) we focus on the individual contributions $\ell_u(\omega; \theta, p)$ to the transformed Whittle likelihood, such that $\ell_u(\theta, p) = \sum_\omega \ell_u(\omega; \theta, p)$. Estimation is carried out by minimizing the discrepancy between the sample spectrum and the theoretical spectral model. This is captured by the second additive term in (7), $\frac{[I_n(\omega)]^p}{[2\pi f(\omega;\theta)]^p}$. There are three possible scenarios.

- The observed spectrum matches the theoretical spectrum. Then, assuming for simplicity that $I(\omega) = f(\omega; \theta) = 1$, $\ell_u(\omega; \theta, p)$ is maximized at $\widehat{p} = 1$. This corresponds to the traditional Whittle estimation, which equally weights the individual likeli-

hood contributions at each frequency $\omega$.

- The periodogram ordinate is much larger than the assumed spectral density, $I(\omega) \gg f(\omega;\theta)$. Then, the numerator of the second additive term in (7), $\left[\frac{I_n(\omega)}{f(\omega;\theta)}\right]^p$ increases fast with $p$, at an exponential rate. At the same time, the denominator $(2\pi)^{p-1}$ also increases with $p$, but at a slower rate, so that the whole fraction is minimized at $\widehat{p} > 1$. Consequently, this implies that the transformed likelihood emphasizes large periodogram observations more carefully than the traditional Whittle likelihood would, suggesting that, based on the specific empirical evidence, modelling spectral peaks is important.

- The sample spectrum at $\omega$ is smaller than the one implied by the assumed model. The likelihood $\ell_u(\omega;\theta,p)$ is maximum at $\widehat{p} < 1$, giving more weight to matching small periodogram ordinates, downweighting large ones. This tells us that the estimator is recognizing spectral peak as outliers and not descriptive of the true data generating process, providing robust estimation. Hence, the estimator is prioritizing more stable spectral features.

## 5.2 ARFIMA Example

Consider now one of the most commonly used long memory models, the autoregressive fractionally integrated moving average (ARFIMA) model. In particular, the spectral density of a stationary ARFIMA(1,$d$,0) model is $f(\omega;\theta) = \frac{\sigma^2}{2\pi}\frac{[2\sin(\omega/2)]^{-2d}}{[1-2\phi\cos(\omega)+\phi^2]} = \frac{\sigma^2}{2\pi}\phi(\omega)\beta(\omega)$, where we assume $-1 < \phi < 1, 0 < d < 0.5$. In presence of long memory ($d \neq 0$), the spectral density is high at low frequencies and it diverges near the zero frequency. In equation (7) we have:

$$\frac{I(\omega)^p}{\sigma^2/(2\pi)\phi(\omega)^p\beta(\omega)^p},$$

measuring the discrepancy between the data and the model. If the data show more low-frequency power (through periodogram peaks) than the spectral model, then the

likelihood will favour increasing $p > 1$ to assign more weight to the near-zero and other frequency regions where such mismatches occur. On the other hand, if the periodogram is flatter than the assumed model, the estimated $p$ will be smaller than one, providing robust model estimation, suggesting to downweight spurious peaks.

Note that the roles of the parameters $d$ and $p$ differ substantially. The memory parameter $d$ controls the persistency attributes of the series, described by the spectral density near zero. On the other hand, $p$ controls the degree of emphasis the likelihood assigns to different parts of the spectrum. Hence, while $d$ is a model parameter, $p$ operates on the loss function used for estimation, emphasizing parts of the data that differ from the model.

# 6 Numerical studies

In this Section, we use simulated data to show how the features of salient time series, outlined in the previous Sections, affect the power parameter $p$. The selected $p$ describes the spectral model best fitting the data in terms of the weighted Whittle likelihood criterion in (7). Specifically, we show that when the time series exhibits salient characteristics, like sudden behavioral changes (corresponding to parameter changes) or repeated increments (or decrements) the optimal value of $p$ will be larger than unity, underscoring the importance of such attributes, and suggesting they should be given more weight in the model fitting procedure. On the other hand, selected values $p \leq 1$ reflect less salient data, providing robustness against outlying periodogram peaks. Given the illustrative purposes of this simulation experiment, we consider optimal selection of $p$ over a specified range of values.

## 6.1 No-salience experiment

First, we illustrate the results of a baseline simulation experiment which allows us to show evidence of no over-estimation of the parameter $p$ capturing the salience properties

outlined above. This consists in generating $R = 1000$ samples of length $T = 3000$ from a Gaussian White Noise process, $x_t \sim WN(0, 1)$, characterized by a constant spectrum over $\omega \in (-\pi, \pi)$, $f(\omega) = \frac{1}{2\pi}$. Following Proietti and Luati (2015), we solve a generalised Yule-Walker system of equations using the non-parametric estimates of the GACV function to obtain the coefficients of the AR approximation of the power process $u_{pt}$. This is done for each value of $p$ over the selected grid. In the present example, for illustrative purposes, we consider positive $p$ in $[0.1, 2]$. The spectral model for the original process $x_t$ is obtained applying the inverse power transform in (4) where estimates of the model parameters have been plugged in. The optimal power parameter is selected minimizing a measure of deviance equal to minus twice the Whittle likelihood, $dev(p) = \sum_j \left[\frac{I(\omega_j)}{\widehat{f}_p(\omega_j)} - \ln \widehat{f}_p(\omega_j)\right]$. Although the theoretical spectrum of the process is flat, the raw periodogram computed on finite samples can show spurious peaks due to sampling variability. Moreover, it does not provide a consistent estimator of the spectral density, and could distort optimal selection of the power $p$ based on the deviance measure $dev(p)$, which contrasts the sample and theoretical spectra for given $p$ values. For these reasons it is preferable and common practice to use a smoothed version of the periodogram, which yields consistent estimation. Let $I_s(\omega)$ be a smoothed periodogram. Here we apply a Daniell kernel to the raw periodogram, as

$$I_s(\omega_j) = \sum_{k=-b}^{b} w_k I(\omega_j + k)$$

, where $w_k$ are the kernel weights such that $\sum_k w_k = 1$, and $h = 2b+1$ is the kernel span. Using a smoothed version of the periodogram lowers its variability and prevents from overestimating the salience features. While the raw periodogram can be characterized by peaks due to random variability, which may influence the selection of $p$ upward, the smoothed periodogram tends to show large ordinates if they are structural, while averaging out spurious peaks. This could rather induce conservative selection of lower $p$ values.

We used a single Daniell kernel with equal weights $w_k = \frac{1}{h}$. The mean selected value over the $R$ replicates is $\widehat{p} = 1.003$. This confirms that the generated data do not show evidence of salient features that could lead the forecaster observing the time series to upweight

certain characteristics and downweight others, in such a way to influence her prediction.

## 6.2 Increasing salience experiments

Here we show the effects of increasing salience levels on the optimal selection of the $p$ parameter. The illustrative Monte Carlo experiments consist in generations of time series of increasing length from different specifications of two classes of models. The first class of models reflects the characteristics of the surprise feature of salience, while the second class corresponds to prominent time series.

To mimic "surprising" time series we kept prominence low and varied the surprise level, so that we can separate the two aspects of salience and analyze only the effects of surprise. This is done by tuning the parameters of the data generating process (DGP), obtained as the sum of two components, specified as follows:

$$\begin{aligned} x_t &= s_t + \varepsilon_t, \\ s_t &= \phi_1 s_{t-1} + \phi_2 s_{t-2} + \eta_t, \\ \eta_t &\sim N(0, \sigma_\eta^2), \\ \varepsilon_t &\sim N(0, \sigma_\varepsilon^2). \end{aligned} \tag{14}$$

In the equation above, $s_t$ represents a cyclical component, generated from the AR(2) model with complex roots, with $\phi_1 = 2r\cos\omega_0$ and $\phi_2 = -r^2$, where $r$ is the damping parameter, determining the persistence of the cycle, and $\omega_0$ is the cycle frequency. We keep persistency low in this first set of Monte Carlo experiments by setting $r = 0.4$ so that the surprise effects are isolated. The noise component $\varepsilon_t$ allows to regulate the strength of the cycle and make it more or less pronounced. Under this design, surprise enters the series through changes in the levels, which manifest as peaks in the spectrum at the cycle frequency $\omega_0$, not necessarily at low frequency. Such changes can be more or less visible according to the magnitude of the noise around them, controlled by the noise variance, $\sigma_\varepsilon^2$.

The second class of models we consider is the class of autoregressive fractionally

integrated moving-average (ARFIMA) models, specifically the ARFIMA(0,d,0), defined as:

$$(1 - B)^d x_t = \varepsilon_t. \tag{15}$$

This is a popular class of long memory models, which allows us to introduce prominence in the data via the memory parameter $d$. The long memory behavior of the series translates in the spectral domain into a pole of $f(\omega)$ near the zero frequency, which shows up in the sample spectrum, $I(\omega)$. The deviance criterion used to select the optimal value of $p$, which tells us about the salience level as we previously defined it, is then minimized for $p$ emphasizing frequencies at which the discrepancy between sample and theoretical spectrum is larger. When salience is present and manifests via either surprise or prominence characteristics, the spectrum shows pronounced peaks at some frequencies (specifically, at low frequencies for prominent data, and in correspondence of the cycle frequency $\omega_0$ in presence of surprise features). This leads the criterion to choose $p > 1$ upweighting these peaks for salient time series.

| **Surprise Level** | **Parameters model (14)** | $\overline{\widehat{p}}$ |
|---|---|---|
| Low | $r = 0.40$, $\sigma_\eta = 0.20$, $\sigma_\varepsilon = 1.00$ | 1.03 |
| Medium | $r = 0.40$, $\sigma_\eta = 0.60$, $\sigma_\varepsilon = 0.80$ | 1.42 |
| High | $r = 0.40$, $\sigma_\eta = 1.20$, $\sigma_\varepsilon = 0.40$ | 1.34 |

Table 1: Average selected $\widehat{p}$ for three levels of surprise. Frequency $\omega_0 = 0.3\pi$ fixed, persistence fixed ($r = 0.40$), varying spectral contrast via $\sigma_\eta$ and $\sigma_\varepsilon$.

| **Prominence Level** | **Parameters model (15)** | $\overline{\widehat{p}}$ |
|---|---|---|
| Low | $d = 0.10$ | 1.37 |
| Medium | $d = 0.25$ | 1.78 |
| High | $d = 0.40$ | 1.61 |

Table 2: Average selected $\widehat{p}$ for three levels of prominence. Surprise kept constant; prominence manipulated through long memory parameter $d$ in ARFIMA$(0, d, 0)$.

Table 1 shows the average optimal $p$ selected by the Whittle deviance criterion over

$R = 1000$ replicates of samples of length $n = 3000$ generated from model (14) with three different specifications of the model parameters, corresponding to different surprise levels. For all the surprise levels we set $r = 0.4$ to keep persistency (related to prominence) low. We selected three values for the standard deviation of the AR(2) innovations, $\sigma_\eta \in \{0.2, 0.6, 1.2\}$, which controls the amplitude of the cyclical component. The strength of the cycle increases as $\sigma_\eta$ increases. The standard deviation of the noise term, $\sigma_\varepsilon$, allows to mitigate the signal strength. The smaller $\sigma_\varepsilon$, the more the cycle dominates the noise. Hence, starting from the combination $\sigma_\eta = 0.2, \sigma_\varepsilon = 1.0$ corresponding to "low surprise", increasing the signal-to-noise ratio $\frac{\sigma_\eta}{\sigma_\varepsilon}$ increases surprise. We can see that the average optimal $p$ is slightly above 1, since we introduced some degree of salience in the data. Moreover, the average optimal $p$ shows a neat increase above one when considering increased surprise levels, consistently with theory, although this path is not monotonic when passing from medium to high surprise level. This is attributable to Monte Carlo variability and to the interpretation of the parameter $p$ as indicator of the presence of salient characteristics rather than a precise measure quantifying the intensity of such features.

To introduce prominence characteristics in the time series, we consider fractional integration through the fractional differencing parameter $d$ and simulate from the ARFIMA(0,d,0) model in (15), letting $d$ vary in $(0, 0.5)$ to produce stationary, long-range dependent observations. As $d$ increases, persistency increases, strengthening the prominence features of the data and their salience. Results in Table 2 show higher $p$ values than those reported in Table 1. This might indicate that prominence is a stronger component of salience than surprise is, and hence it has a deeper impact on individuals' expectations. Similarly to the previous experiments with surprise characteristics, $\overline{\widehat{p}}$ shows a sharp increase from low to medium prominence levels, for each given sample size $n$. The effects of further increasing persistency of the data from medium to high are marginal and not always correspond to a monotonic increase in the optimal $p$. Overall, these results corroborate our theoretical findings by showing that the power parameter $p$ defining the class of spectral models (4) and selected for a given time series realization by optimizing the weighted

Whittle deviance criterion, does not exceed unity in absence of salience, and lies above 1 for salient data, detecting stronger effects of prominence than surprise.

In the following we provide a graphical illustration of the theoretical results and the instruments used to detect salience. We used different models from the previous ones in (14) and (15) to simulate data characterized by a given level of salience based on surprise and prominence features. Let us start by a low salience example: in $R$ we generated a time series of length $T = 2000$ from a zero-mean ARMA(1,1) model with AR and MA parameters respectively $\phi = 0.3$ and $\theta = 0.5$.

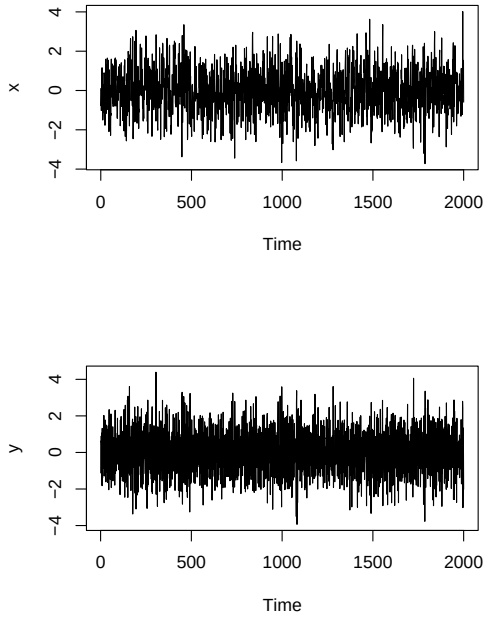



Figure 7: $x_t$: ARMA(1,1); $y_t$: first differenced series.

The series shown in figure 7 along with its first differences, is stationary and short-memory over the entire sample, hence its first- and second-order properties do not change over time. This implies that surprise effects are minimal.

Moreover, the simulated series is characterized by short-memory and zero constant mean. This means that changes of a given sign are likely followed by changes of the opposite sign, making the prominence of the series very low.

By using the sample estimates of the GACV we solved the generalised Yule-Walker system of equations and obtained spectral estimates $\widehat{f}_p$ corresponding to a given $p$ in a grid

between $-3$ and 3. The optimal spectral estimate is selected by minimising the deviance measure with respect to $p$. The optimal negative value $p = -1$ amounts to fitting a MA model to the data.

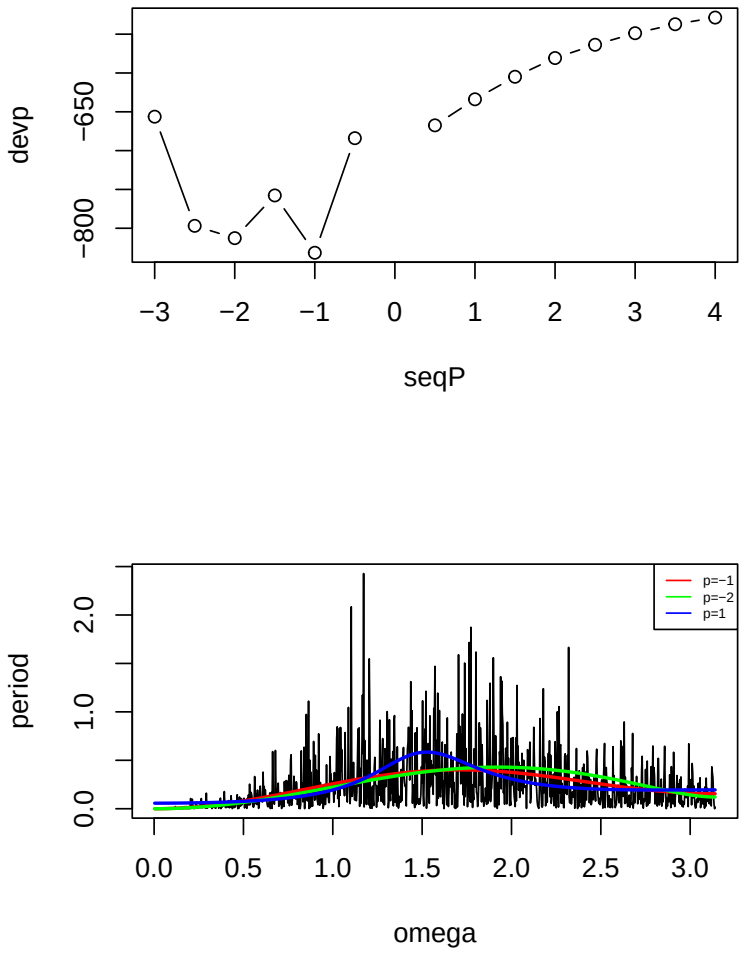



Figure 8: Top: deviance measure vs $p$. Bottom: periodogram estimate and spectral estimates $\widehat{f}_p$ with $p = -2, -1, 1$.

From figure 8 it can be seen that negative values of $p$ are better suited than positive ones to describe the spectrum of the current series. The spectral estimate corresponding to the optimal $p = -1$, shown in the bottom panel, appears to be the most flat among the three, with a smaller range compared to $\widehat{f}_p$ with $p = -2, 1$. The green curve corresponding to $p = -2$ associated to estimation of the GACV emphasizes troughs in the periodogram, as negative $p$ assign higher weight to lower periodogram ordinates, while the curve corre-

sponding to $p = 1$ equally weights all periodogram ordinates. Neither $p = 1$ nor $p = -2$ reflects the second order properties of the current series in the frequency domain, but the flatter curve defined by $p = -1$ is more suitable.

We proceed in our analysis on simulated data by introducing in the time series one additional element increasing its salience. The second series is generated by simulating the first $T_1 = 1000$ observations from an AR(1) model with $\phi = 0.7$ and the last $T_2 = T_1$ observations from a MA(1) model with $\theta = 0.3$. This increases the salience of the series through surprise effects caused by a behavioral change over time.

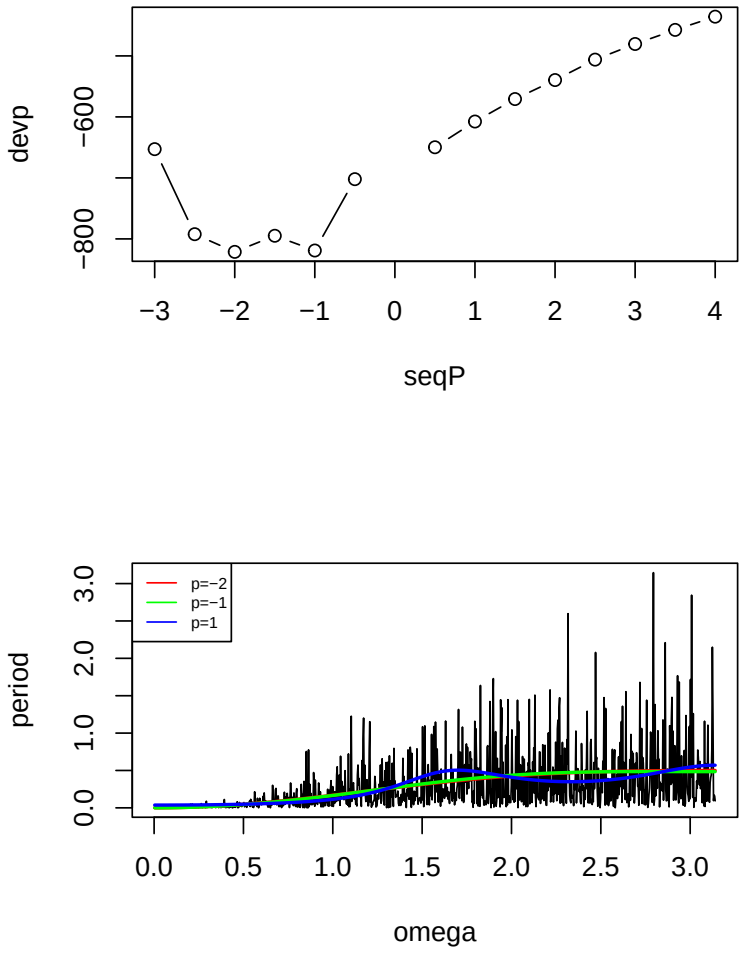



Figure 9: Top: deviance measure vs $p$. Bottom: periodogram estimate and spectral estimates $\widehat{f}_p$ with $p = -2, -1, 1$.

The selected spectral model for this series corresponds to $p = -2$. The spectral density

function is shown in figure 9 and almost entirely overlaps with that for $p = -1$, being slightly above it for higher periodogram ordinates and slightly below it for lower ones.

In the next simulated series we introduce long memory behavior. We simulate the differenced series from a stationary ARFIMA(1,d,1) model with the same AR and MA parameters as the first simulated series, and with memory parameter $d = 0.3$. The levels of the series are recovered by cumulative sums, and show nonstationary long memory behavior. Both series are plotted in figure 10. The goodness-of-fit criterion based on deviance is minimum for $p = 2$, as shown in Figure 11. The red curve ($p = 2$) captures the long memory properties of the series, highlighting the presence of a spectral pole at zero.

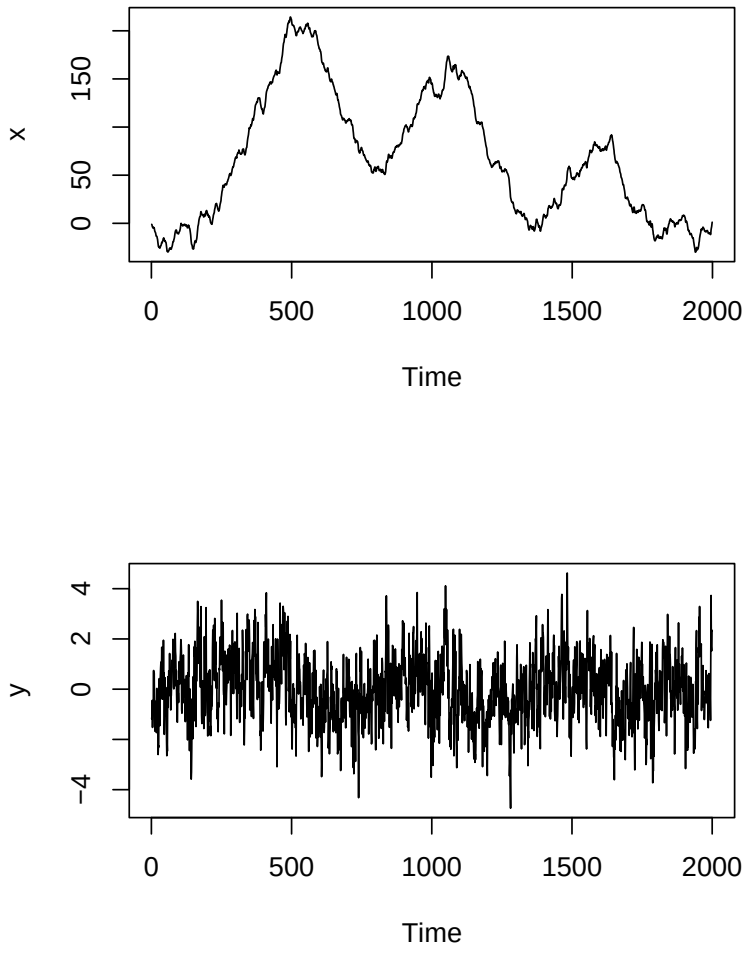



Figure 10: $x_t$: ARFIMA(1,d,1); $y_t$: first differenced series.

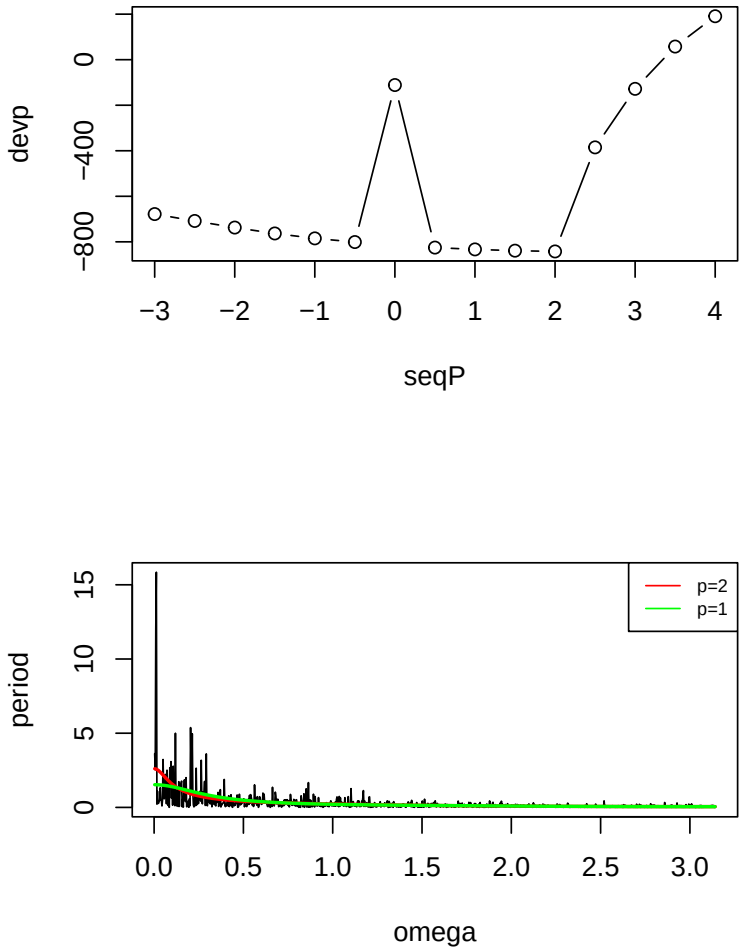



Figure 11: ARFIMA(1,d,1) plot. Top: deviance measure vs $p$. Bottom: periodogram and spectral estimates $\widehat{f}_p$ with $p = 2, 1$.

Finally, we consider a stationary long memory series with a change in the Wold parameter, by simulating the first $T_1 = 1000$ observations of the $x_t$ series from an ARFIMA(1,d,0) model with $\phi = 0.7$ and $d = 0.45$ and the last $T_2 = T_1$ observations from an ARFIMA(0,d,1) with $\theta = 0.2$ and $d = 0.45$. The levels and first differences of the series are plotted in Figure 12.

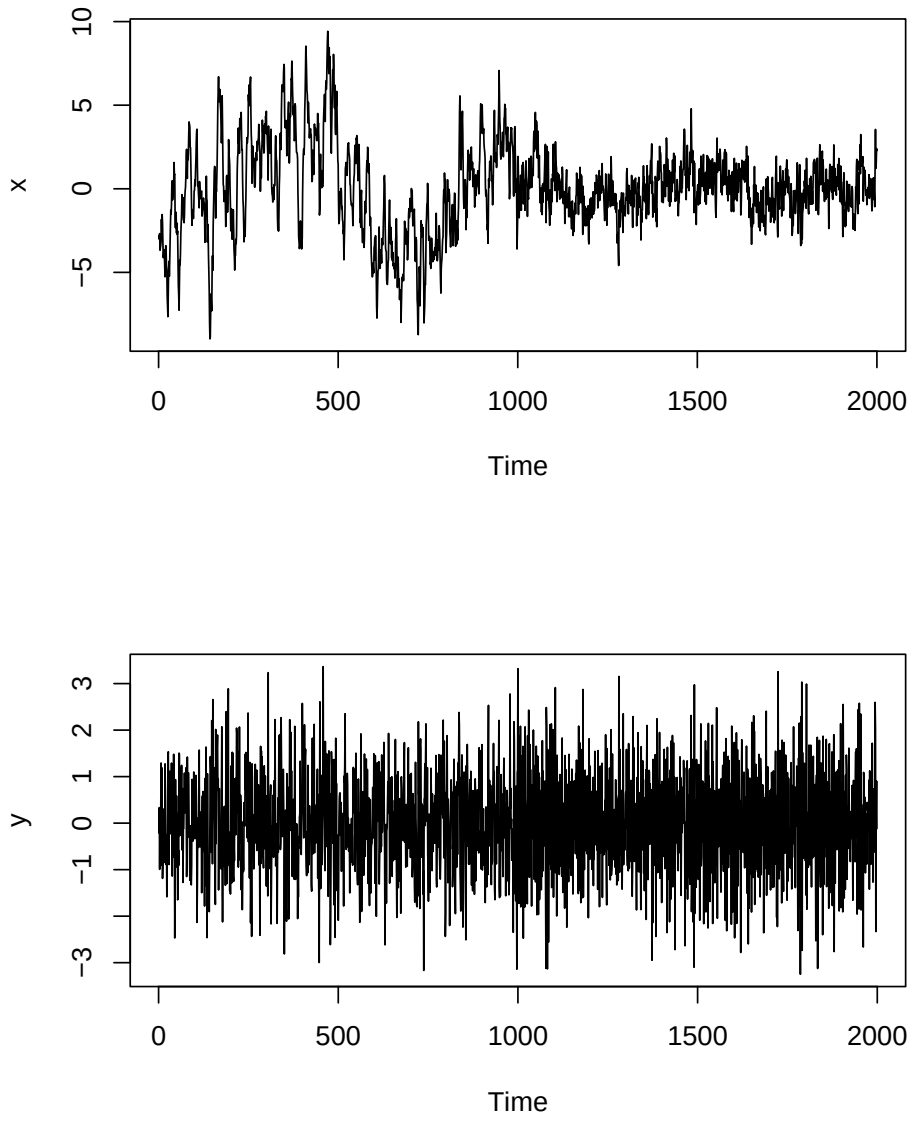



Figure 12: $x_t$: ARFIMA(1,d,1) with parameter changes; $y_t$: first differenced series.

The selected spectral model for this series, based on the deviance measure in Figure 13 corresponds to $p = -2.5$. The bottom panel of the series plots the spectral estimates computed on the differenced series and shows that at the minimum periodogram ordinate, at zero frequency, $\widehat{f}_p$ with $p = -2.5$ lies below the other curves, while the spectral estimate for $p = -1.5$ is slightly above it. The curve corresponding to $p = 1.5$ is the most flat one among the three compared. While $\widehat{f}_{-1.5}$ and $\widehat{f}_{-2}$ partially overlap.

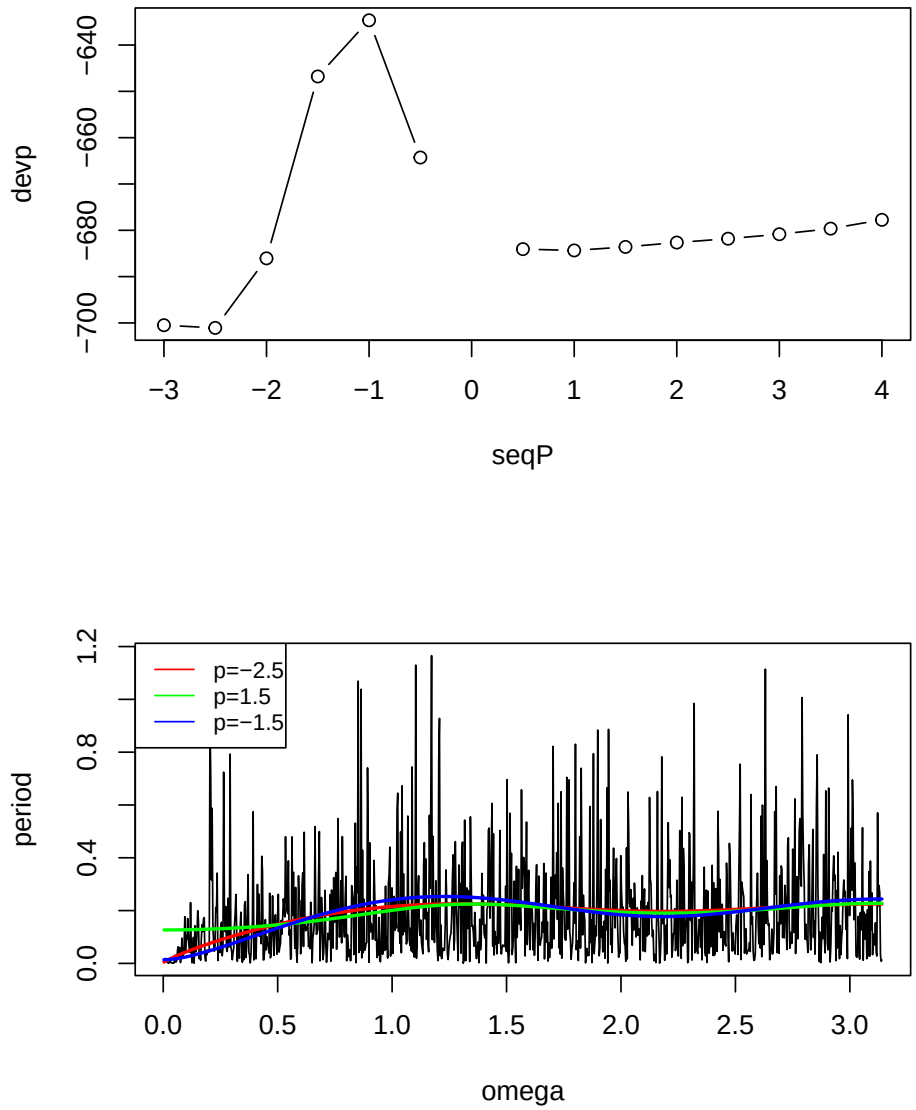



Figure 13: ARFIMA(1,d,1) with parameter changes. Top: deviance measure vs p. Bottom: periodogram and spectral estimates $\widehat{f}_p$ with $p = -2.5, -1.5, 1.5$

# 7 Empirical application

This Section applies the estimation procedure described above to U.S. inflation and inflation expectations. The objective is to characterize the temporal dependence structure of aggregate expectation dynamics and compare it to the dynamics of realized inflation. In the framework introduced in the previous Sections, the parameter $p$ measures the rel-

ative importance of certain frequency contributions to the total variation of the process: low values indicate that periodogram peaks are downweighted, suggesting that they are spurious, while high values indicate that most variation is attributable to sinusoidal components at specific frequencies, which receive higher weight in the estimation procedure.

To interpret the estimates economically, the analysis considers three related time series: realized inflation, aggregate inflation expectations, and forecast errors. Together, they allow us to distinguish between salience in the underlying macroeconomic signal, the component incorporated into beliefs, and the residual information not reflected in expectations.

## 7.1 Data

Inflation realizations are constructed from the Consumer Price Index for All Urban Consumers (CPI-U), obtained from the Federal Reserve Bank of St. Louis FRED database (series `CPIAUCSL`). The monthly index is converted into a twelve-month inflation rate in order to match the horizon of survey expectations.

Aggregate inflation expectations are taken from the University of Michigan Surveys of Consumers, using the FRED series `MICH`, which reports one-year-ahead expected inflation at monthly frequency. This series provides an aggregate measure of household inflation beliefs and can be interpreted as a representative indicator of consumer expectations at each point in time.

The Michigan expectations series consists of monthly observations from January 1, 1978 to February 1, 2026. After aligning realizations and expectations in time, the empirical analysis uses a monthly sample starting in January 1978 and ending at the latest date for which twelve-month realized inflation can be matched to the expectation observed at the time the forecast was formed.[1]

Forecast errors are defined as the difference between realized inflation over the following twelve months and the expectation reported at the time the forecast was made. This variable captures the component of inflation not incorporated into beliefs when expecta-

[1] Because realized inflation is defined over the following twelve months, the final year of CPI observations cannot be used directly in the comparison and is therefore excluded after alignment.

tions are formed and therefore represents residual information relative to expectations.

After aligning realizations and expectations in time, the final dataset consists of three monthly time series:

- realized inflation,
- expected inflation (Michigan, one-year ahead),
- forecast errors.

Within the empirical framework of the paper, realized inflation represents the underlying macroeconomic signal with potentially salient information, expectations capture perceived inflation dynamics, and forecast errors measure the remaining unincorporated information.

## 7.2 Results

For each of the three series — realized inflation, aggregate expectations, and forecast errors — the optimal selection procedure described in Section 4 is applied to first differenced data, inducing stationarity. The optimal selection of $p$ is conducted over a grid of positive values for interpretability purposes, with $p \in [0.01, 10]$. The analysis is conducted separately for the three variables in order to compare the temporal dependence structure of their dynamics. The optimal parameter $p$ is selected using the frequency-domain criterion and inference is obtained via bootstrap. In this framework, the parameter $p$ selects the optimal model in the class of spectral models (4) and may highlight salient features of the spectrum. Low values of $p$ correspond to a relatively even weighting of spectral contributions in the Whittle-type criterion, while high values indicate that the fit is driven more strongly by a subset of large spectral components.

Table 3 reports the estimated values of $p$, together with bootstrap standard errors obtained over $B = 700$ bootstrap replications.

The results reported in Table 3 suggest that the salience parameter is above unity across the three series considered. For realized inflation, the estimated value is rela-

Table 3: Estimated power parameter $p$

| Series | $\widehat{p}$ | Std. Error |
|---|---|---|
| Realized inflation | 9.58 | 1.73 |
| Inflation expectations | 4.59 | 2.03 |
| Forecast errors | 3.66 | 1.08 |

tively large, with $\widehat{p} = 9.58$, indicating a pronounced departure from the neutral benchmark $p = 1$. This frequency-domain evidence is consistent with the presence of major macroeconomic developments affecting inflation over the sample period, including the high-inflation episodes of the late 1970s and early 1980s as well as the post-2021 inflation surge. However, the parameter $p$ itself does not identify the timing or nature of such developments; rather, it summarizes the extent to which variation is concentrated in salient dynamic components.

Inflation expectations also display a value of $p$ clearly above one, with $\widehat{p} = 4.59$, although lower than that of realized inflation. This indicates that expectations exhibit substantial spectral structure, but with a more diffuse distribution of variance across frequencies than realized inflation. In other words, expectations inherit part of the systematic dynamics of inflation, yet their variability is less concentrated in a narrow set of oscillatory components, consistent with a partial filtering or smoothing of the underlying inflation process.

This pattern is consistent with the idea that expectation formation incorporates dynamic structure of realized inflation in an attenuated way. The parameter $p$ does not identify the specific mechanism behind this attenuation, yet it suggests that the transmission from inflation to expectations is incomplete and subject to selective incorporation of its dynamic components.

At the same time, the bootstrap standard errors are non-negligible relative to the point estimates, especially for expectations. This reflects the flatness of the deviance criterion used to select $p$ and the resulting sampling variability of the optimal parameter selector. Therefore, the point estimates should be interpreted primarily as indicating the presence and relative strength of salient spectral structure, rather than as sharply

identified structural constants.

In the case of forecast errors, the selected optimal value $\widehat{p} = 3.66$ still lies well above unity, although it is substantially smaller than the values obtained for the other two series. This indicates that forecast errors are not spectrally white. Instead, they retain structured dynamic components, implying that expectations do not eliminate all systematic variation in the inflation process. In other words, the dynamic structure of inflation is only partially internalized by expectations. The salient component of inflation not incorporated into expectations is therefore not purely random noise.

## 7.3 Discussion

The empirical findings can be interpreted in relation to models of salience-based expectation formation. In the salience framework, agents do not process all available information symmetrically; instead, they place disproportionate weight on the most prominent or salient aspects of the economic environment. As a consequence, expectations may reflect certain components of macroeconomic dynamics more strongly than others. From the perspective of salience theory, this pattern is consistent with the idea that agents selectively incorporate the most prominent features of macroeconomic fluctuations while other components receive less attention.

Importantly, this interpretation should not be viewed as a rejection of rational expectations. Rational expectations allow forecast errors to arise from unpredictable structural changes in the economic environment. Rather, the results are consistent with a broader class of models in which expectation formation reflects partial or selective processing of macroeconomic information.

Specifically, the evidence points to an incomplete transmission of inflation's dynamic structure into expectations. Although beliefs internalize part of the systematic variation in realized inflation, they do not fully eliminate its structured components. As a result, residual forecast errors retain some dynamic structure rather than behaving as purely diffuse disturbances.

Overall, the dynamic structure of realized inflation attributable to salient events is

only partially reflected in expectations. While beliefs incorporate a substantial share of the systematic variation present in inflation, they do not reproduce its full spectral structure. Such results can also be interpreted in light of models of salience and diagnostic expectations. In these frameworks, agents do not process all available information symmetrically; instead, they overweight the most prominent aspects of the economic environment while placing relatively less weight on less visible components of macroeconomic dynamics.

# 8 Conclusion

This paper proposes a statistical framework to detect salience and its effects in time series. By focusing on the dimensions of surprise and prominence, and by embedding them in the spectral representation of a stochastic process, the paper provides tools to study how salient features of economic dynamics may shape expectation formation.

The main contribution is to provide a statistical characterization of salience for temporally dependent data, thereby extending the salience framework from cross-sectional settings to time series. Second, it develops an operational strategy that can be applied to observed historical data. Third, it shows how the salience parameter can be used to study expectation formation in macroeconomics. This parameter is selected optimizing a goodness-of-fit Whittle-type criterion. It summarizes the extent to which emphasizing spectral peaks improves the model fit, indicating evidence of salience and its effects. In this framework, salience is not defined in terms of isolated large observations, but in terms of the extent to which the dynamic structure of a series is dominated by systematic components that stand out relative to the rest of the process.

The numerical exercises show that the proposed approach is able to distinguish between processes with little salience and processes whose dynamics are strongly structured. The empirical application to U.S. inflation and inflation expectations illustrates the relevance of the method in a macroeconomic context. The estimates indicate that realized inflation exhibits the strongest spectral structure, aggregate expectations also display

substantial structure, and forecast errors remain non-diffuse. This suggests that expectations reflect part of the systematic variation in inflation, but do not fully incorporate its dynamic structure.

From an economic perspective, the results are consistent with the view that expectation formation is not based on a complete and symmetric processing of all available macroeconomic information. Instead, some components of inflation dynamics appear to be internalized more strongly than others. This interpretation is broadly consistent with models of salience, diagnostic expectations, limited attention, and infrequent updating, although the evidence should not be read as a direct rejection of rational expectations. Rather, the findings suggest that the transmission from realized inflation to expectations is incomplete, leaving residual structured dynamics in forecast errors.

More broadly, the paper suggests that salience can be studied statistically in observed macroeconomic data and not only inferred from experimental or survey evidence. This opens the possibility of bringing behavioral notions of attention into formal time series analysis, thereby creating a bridge between behavioral economics and empirical macroeconometrics.

# References

Balcombe, K., Fraser, I., and Sharma, A. (2021). An econometric analysis of salience theory. *Bulletin of Economic Research*, 73(4):545–554.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly journal of economics*, 127(3):1243–1285.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and asset prices. *American Economic Review*, 103(3):623–628.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2019). Memory and reference prices: An application to rental choice. In *AEA Papers and Proceedings*, volume 109, pages 572–576. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020). Memory, attention, and choice. *The Quarterly journal of economics*, 135(3):1399–1442.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2022). Salience. *Annual Review of Economics*, 14(1):521–544.

Briand, E., Marcellino, M., and Stevanović, D. (2025). *Inflation, Attention and Expectations*. CIRANO.

Ilut, C. and Valchev, R. (2023). Economic agents as imperfect problem solvers. *The Quarterly Journal of Economics*, 138(1):313–362.

Luati, A., Proietti, T., and Reale, M. (2012). The variance profile. *Journal of the American Statistical Association*, 107(498):607–621.

Miao, J., Wu, J., and Young, E. R. (2022). Multivariate rational inattention. *Econometrica*, 90(2):907–945.

Parsley, D. and Popper, H. (2024). Climate change salience and international equity returns. *Journal of Economic Behavior & Organization*, 226:106685.

Proietti, T. and Luati, A. (2015). The generalised autocovariance function. *Journal of Econometrics*, 186(1):245–257.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690.

Van Der Vaart, A. W. and Wellner, J. A. (1996). Weak convergence. In *Weak convergence and empirical processes: with applications to statistics*, pages 16–28. Springer.

Whittle, P. (1953). Estimation and information in stationary time series. *Arkiv för matematik*, 2(5):423–434.

Zanetti Chini, E. (2023). Can we estimate macroforecasters' mis-behavior? *Journal of Economic Dynamics and Control*, 149:104632.