XVII Congreso Galego de Estatística e Investigación de Operacións
Ourense, 23 - 25 de outubro de 2025

Addressing Structural Breaks in the Seasonal Adjustment of Official Statistics

Cheyenne Amoroso[1], Siem Jan Koopman[2], Carolina García-Martos[3], Germán Aneiros[1], José A. Vilar[1], Manuel Oviedo de la Fuente[1] and Mario Francisco-Fernández[1].

[1] Grupo MODES, CITIC, Universidade da Coruña, España.
[2] Department of Econometrics, VU University Amsterdam, Netherlands.
[3] Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, España.

ABSTRACT

Seasonal adjustment is a key task in official statistics to better understand the dynamics of time series. Long socio-economic series, however, are often affected by structural breaks due to crises or institutional reforms. Ignoring such breaks leads to distorted adjustments and misleading interpretations. This paper proposes a state space approach based on the Basic Structural Model with time-varying disturbance variances, allowing for both abrupt and gradual changes. Breakpoints are detected through standardized innovations and smoothed residuals, serving as empirical measures of local volatility. The method ensures a coherent decomposition while remaining fully compatible with the Kalman filter, making it operationally feasible for official statistics. We illustrate the framework on Spanish Statistical Office (INE) time series, showing improved coherence of the component decomposition around identified breaks.

Keywords: Seasonal adjustment; Structural breaks; State space models; Official Statistics.

1. INTRODUCTION

A central task in official statistics is the production of seasonally adjusted series, that is, series from which recurrent and systematic fluctuations, such as those caused by holidays or weather conditions, have been removed. This adjustment is crucial, as seasonal patterns can obscure both short- and long-term movements, thereby hindering the identification of underlying dynamics. The resulting adjusted series provides a clearer representation of the fundamental behaviour of the variable, enhancing interpretability and supporting more reliable forecasting and evidence-based decision-making.

With the aim of harmonising practices across Europe and improving the comparability of national infra-annual statistics, the European Statistical System (ESS) published in 2009 the first set of *Guidelines on Seasonal Adjustment*, Eurostat (2009). A revised edition was released in 2015, Eurostat (2015), presenting both theoretical foundations and practical recommendations in an accessible framework. Following these initiatives, the Spanish Statistical Office (INE) issued its own guidelines for the treatment of socio-economic time series, INE (2019). These efforts have contributed to the standardisation of seasonal adjustment procedures within official statistics, while also encouraging methodological innovation.

Within the ESS framework, two approaches are recognised as equally valid for seasonal adjustment: ARIMA model-based adjustment and fixed filter-based adjustment. Currently, the Spanish Statistical Office applies the well-established `TRAMO-SEATS` method, Gómez and Maravall (1996), an ARIMA model-based approach which involves two main stages: first, a Reg-ARIMA model is estimated to identify and correct outliers while accounting for calendar effects such as holidays or variations in month length. Then, signal extraction is performed using Wiener–Kolmogorov filtering, a sophisticated technique that separates the underlying signal from irregular noise.

However, socio-economic time series often span long periods during which economic conditions, social dynamics, and institutional frameworks may change substantially, leading to shifts in their underlying patterns. Such developments challenge the stability assumption underlying traditional seasonal adjustment procedures, which are typically based on a single parametric model fitted to the entire sample. A well-documented example is the 2008 financial crisis, which introduced profound and lasting structural changes. In this context, two issues become particularly relevant for National Statistical Offices: the identification of the points in time at which structural breaks occur and the development of a robust methodology for seasonal adjustment that remains reliable in their presence.

The present work addresses these challenges by proposing a new framework based on state space models, which provides the necessary flexibility to explicitly account for structural breaks and accommodate time-varying dynamics. This ensures a coherent decomposition of trend, seasonal, and irregular components in long series affected by instability, making the methodology particularly well-suited for official statistics.

The remainder of this paper is structured as follows. Section 2 introduces the state space modelling framework, highlighting its relevance for seasonal adjustment. Section 3 presents time-dependent models as a natural extension to handle changing dynamics. Section 4 discusses the detection of structural breakpoints and their implications for time series analysis. Section 5 illustrates the proposed methodology through an application to real socio-economic series produced by the Spanish Statistical Office. Finally, Section 6 summarises the main findings and outlines directions for future research.

2. STATE SPACE MODELS

State space models provide a general and flexible framework for modelling time series, allowing a wide range of dynamic processes to be represented through unobserved components that evolve over time. This makes them particularly well suited for tasks such as seasonal adjustment and trend-cycle decomposition. In order to establish the foundation for the proposed methodology, this section reviews the fundamental concepts and notation of state space models, as outlined by Durbin and Koopman (2012).

In the general linear Gaussian state space model, the observed series y_t is linked to a latent state vector α_t through two sets of equations:

$$y_t = Z_t\alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, H_t), \tag{1}$$
$$\alpha_{t+1} = T_t\alpha_t + R_t\eta_t, \qquad \eta_t \sim \mathcal{N}(0, Q_t). \tag{2}$$

Equation (1), known as observation equation, relates the observed variable y_t to the latent state α_t via the design matrix Z_t. The disturbance term ε_t accounts for measurement noise or other short-term effects not captured by the state. Equation (2), referred to as state equation, describes the evolution of α_t over time via the transition matrix T_t, with innovations η_t having covariance matrix Q_t. The matrix R_t allows these innovations to influence specific elements of the state vector rather than the entire vector.

A particularly relevant specification within the state space framework is the *Basic Structural Model* (BSM), which decomposes the series into interpretable components

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \tag{3}$$
$$\mu_{t+1} = \mu_t + \beta_t + \eta_{\mu,t}, \tag{4}$$
$$\beta_{t+1} = \beta_t + \eta_{\beta,t}, \tag{5}$$
$$\gamma_{t+1} = -\sum_{j=1}^{s-1} \gamma_{t+1-j} + \eta_{\gamma,t}, \tag{6}$$

where μ_t is the level, β_t the slope, γ_t the seasonal component of period s, and ε_t the irregular disturbance. The innovations $\eta_{\mu,t}, \eta_{\beta,t}, \eta_{\gamma,t}$ are mutually uncorrelated white noise processes with variances $\sigma^2_\mu, \sigma^2_\beta, \sigma^2_\gamma$, respectively.

The BSM offers several advantages for seasonal adjustment. Its additive structure provides a transparent interpretation of the decomposition, while the stochastic treatment of components allows for flexibility in adapting to changes in the data. It is also the foundation of methods currently used in official statistics, such as `TRAMO-SEATS`, which makes it a natural benchmark for methodological extensions.

In its standard formulation, the BSM assumes constant variances for the state disturbances. While convenient, this assumption may be unrealistic for long socio-economic series affected by events such as economic crises, institutional reforms, or methodological changes. In our context, relaxing this assumption is key: by allowing the variances of the disturbances to change across regimes, the BSM becomes a powerful tool to capture structural breaks without losing coherence in the decomposition. This formulation not only preserves interpretability and flexibility, but also provides the natural statistical setting, through the Kalman filter and its extensions, to accommodate structural instability in official statistics.

3. PIECEWISE DYNAMIC MODEL

Building on the state space formulation of the Basic Structural Model, we now turn to the problem of seasonal adjustment in the presence of structural breaks. Such breaks challenge the assumption of parameter constancy that underpins traditional seasonal adjustment methods, as a single model estimated over the full sample may fail to capture the heterogeneity between stable and unstable periods. A natural way to address this issue within the state space framework is to allow for regime-specific dynamics while maintaining a unified structure linking the regimes. Rather than treating each sub-period in isolation, our approach models the transition across regimes in a way that preserves both the interpretability and coherence of the extracted components.

In this section, we focus on the case where the locations of structural breaks are known in advance and we denote them as a and b. For expositional clarity, we distinguish three regimes: a pre-break period or Regime I ($a \leq t$), a transitional phase or Regime II ($a < t < b$) and a post-break period or Regime III ($t \geq b$). Regimes I and III are assumed to reflect relatively stable dynamics, while Regime II corresponds to the period surrounding the break, where parameter instability is expected. The key element of the proposed approach lies in the treatment of the disturbance variances across these regimes. For Regimes I and III, independent estimates of the BSM are obtained via maximum likelihood, yielding variance structures that serve as anchors. The transitional period is then modelled by allowing the variances of the state disturbances to evolve between these anchors, ensuring continuity in the decomposition and avoiding spurious dynamics that would arise from fitting an independent model to an unstable interval.

The distinction between abrupt and gradual changes plays a central role in this specification. An *abrupt change* corresponds to a sudden shift in the variance of a disturbance, typically linked to extraordinary events such as financial crises, institutional reforms, or methodological revisions in statistical practice. In contrast, a *gradual change* reflects a progressive adjustment in the variance, arising from slow-moving transformations in the economy or the incremental adoption of new behavioural patterns. Based on this clarification, the treatment of the variances is as follows:

- **Gradual changes:** when the variance of a disturbance changes gradually, it is interpolated smoothly between the pre- and post-break regimes using weights $w_t \in [0, 1]$, which increase from 0 in Regime I to 1 in Regime III:

$$w_t = \begin{cases} 0, & t \leq a, \\ \dfrac{1}{1 + \exp\left(-\dfrac{t - \frac{a+b}{2}}{\frac{b-a}{2} - \left| t - \frac{a+b}{2} \right|} \right)}, & a < t < b, \\ 1, & t \geq b. \end{cases} \tag{7}$$

  The variance is then computed as a weighted average of the pre- and post-break variances:

$$\sigma_t^2 = (1 - w_t)\sigma_{(I)}^2 + w_t \sigma_{(III)}^2.$$

- **Abrupt changes:** When diagnostics indicate a sudden shift in the variance, it is estimated directly within the transition period without interpolation. In this case, the variance is specified piecewise:

$$\sigma_t = \begin{cases} \sigma_{(I)}^2, & t \leq a, \\ \sigma_{(\mathrm{II})}^2, & a < t < b, \\ \sigma_{(III)}^2, & t \geq b. \end{cases} \tag{8}$$

  This specification isolates the abrupt shock while maintaining coherence of the variance outside the transition interval.

The treatment of individual components reflects their different statistical roles. For the slope (β_t) and seasonal (γ_t) components, variances are always assumed to change gradually, and smooth interpolation is applied. By contrast, the irregular (ε_t) and level (μ_t) components may undergo abrupt or gradual changes, which are determined through diagnostic assessment. This distinction ensures that the decomposition adapts flexibly to the nature of the underlying instability.

All unknown quantities in the model are estimated jointly by maximising the Gaussian likelihood via the Kalman filter. The Kalman smoother then produces coherent estimates of the trend, seasonal, and irregular components across regimes, ensuring that the resulting seasonal adjustment remains statistically consistent even in the presence of structural breaks.

4. DETECTION OF STRUCTURAL BREAKPOINTS

Up to this point, the methodology has assumed that the locations of structural breaks are known. However, in practice, breakpoints are often unknown and must be inferred from the data. Accurate detection of these points is crucial, as misidentifying the timing or number of breaks can lead to biased variance estimates and distort the seasonal adjustment. In this section, we describe a systematic approach for detecting structural breaks in long time series, focusing on changes in the variance of the state disturbances, which are central to the regime-based modelling strategy introduced in the previous section.

The detection procedure exploits information from residual diagnostics, smoothed component estimates, and variance patterns. The starting point is to fit a BSM to the full sample under the (incorrect) assumption of parameter constancy. This deliberate misspecification provides residuals that are informative about structural instability, as deviations from white-noise behaviour typically signal potential breakpoints.

The first set of diagnostics is given by the one-step-ahead prediction errors, also known as *innovations* or *model residuals*. These are computed as

$$\nu_t = y_t - \mathbb{E}(y_t \mid y_{1:t-1}),$$

with distribution

$$\nu_t \sim \mathcal{N}(0, F_t), \quad F_t = Z_t P_{t|t-1} Z_t^\top + H_t,$$

where $P_{t|t-1} = \mathrm{Var}(\alpha_t \mid y_{1:t-1})$ is the a priori state prediction error variance. These are the key quantities used in the Kalman filter. For diagnostic purposes, the innovations are standardised using F_t, obtaining the standardised residuals

$$\tilde{\nu}_t = F_t^{-1/2} \nu_t,$$

which, under correct specification, should resemble i.i.d. standard normal noise. Departures from this behaviour provide the first indication of possible structural breaks.

Since innovations measure the unexplained part of the model, their squared values act as local proxies for volatility. If the system is subject to variance breaks, these squared residuals should display persistent shifts away from their expected value. To smooth high-frequency noise and highlight persistent deviations, we compute rolling means of the squared residuals using a symmetric window. For a chosen window size k, we define

- Right-aligned mean:

$$\bar{r}_t^{(R)} = \frac{1}{k} \sum_{j=t-k+1}^{t} r_j,$$

- Left-aligned mean:

$$\bar{r}_t^{(L)} = \frac{1}{k} \sum_{j=t}^{t+k-1} r_j,$$

where $r_t = \tilde{\nu}_t^2$. These rolling means act as local estimators of the variance, with directional sensitivity: the left-aligned mean anticipates variance increases, while the right-aligned mean captures their persistence.

We then set empirical thresholds and identify exceedances as signals of instability. The *lower bound a* of a breakpoint region is derived from $\bar{r}_t^{(L)}$, while the *upper bound b* is obtained from $\bar{r}_t^{(R)}$. The resulting interval $[a, b]$ is interpreted as a *breakpoint region*, consistent with the working hypothesis that changes occur in the variances of the state disturbances.

This methodology is conceptually related to classical change-point detection techniques such as CUSUM or MOSUM. In those methods, cumulative or moving sums of residuals are monitored to detect departures from stability. Here, instead of cumulative means, we use rolling averages of squared residuals and smoothed disturbances, explicitly targeting changes in variance rather than mean shifts. This makes the procedure well-suited to the regime-switching variance framework assumed in Section 3.

Once a structural change window $[a, b]$ has been identified, the next step is to determine the nature of the change: whether it is abrupt or gradual, and to which component it can be attributed. This is done by analysing the *smoothed residuals* (or disturbances), defined as the conditional expectations of the structural shocks given the full sample:

$$\hat{\eta}_t = \mathbb{E}(\eta_t \mid y_{1:T}),$$

where η_t collects the disturbances associated with the level, slope and seasonal components. Under correct specification, the smoothed disturbances are conditionally unbiased estimators of the true shocks, and their squared values behave as empirical proxies of the corresponding variances:

$$\mathbb{E}\big[\hat{\eta}_t^2\big] \approx \mathrm{Var}(\eta_t).$$

To characterise the dynamics of the variance around a breakpoint, we split the time span into three non-overlapping periods: before the break ($t < a$), during the break interval ($a \leq t \leq b$), and after the break ($t > b$). Within each segment we compute the empirical variance estimator

$$\hat{\sigma}^2_{(j)} = \frac{1}{n_j} \sum_{t \in I_j} \hat{\eta}_t^2,$$

where I_j denotes the set of indices belonging to segment $j \in \{\text{pre}, \text{break}, \text{post}\}$ and $n_j = |I_j|$. Thus, $\hat{\sigma}^2_{(j)}$ provides a non-parametric way to approximate the regime-specific variances of the state disturbances.

By comparing the three local variance estimates, we establish heuristic rules to classify the structural change. An *abrupt change* is identified if

$$\hat{\sigma}^2_{\text{break}} \gg \max\{\hat{\sigma}^2_{\text{pre}}, \hat{\sigma}^2_{\text{post}}\},$$

with $\hat{\sigma}^2_{\text{pre}} \approx \hat{\sigma}^2_{\text{post}}$, indicating that the volatility increase is localised in the interval $[a, b]$. Conversely, a *gradual change* is characterised by a monotonic trend across the three segments,

$$\hat{\sigma}^2_{\text{pre}} < \hat{\sigma}^2_{\text{break}} < \hat{\sigma}^2_{\text{post}} \quad \text{(or vice versa)},$$

consistent with a progressive structural transformation in the disturbance variance.

This diagnostic layer links naturally with the piecewise dynamic model: the breakpoint intervals $[a, b]$ guide the specification of Regime II, and the abrupt/gradual classification determines whether variance interpolation or discrete shifts should be applied. In this way, detection and modelling are fully integrated into a coherent seasonal adjustment framework.

5. APPLICATION TO OFFICIAL STATISTICS

To illustrate the practical relevance of the proposed framework, we analyse the monthly Spanish Industrial Production Index (IPI, General index) published by the Spanish Statistical Office (INE). The sample employed in the empirical illustration covers January 1992 to December 2019, a span chosen to provide a long pre- and post-crisis perspective while avoiding the large and atypical disturbances introduced by the COVID-19 pandemic. As one of the core short-term indicators of industrial activity, the IPI plays a central role in monitoring the business cycle, and its seasonal adjustment is of direct importance for economic analysis and policy support.

From an economic standpoint, the IPI is particularly interesting as it reflects a profound structural transformation in the Spanish economy. The 2008 financial crisis marked a turning point: industrial output fell sharply and, unlike in other European economies, the Spanish index never returned to its pre-crisis level, as we can see in Figure 1. This persistent gap cannot be interpreted merely as a cyclical fluctuation but rather as evidence of a deeper reallocation of resources. In particular, the crisis accelerated the decline of construction-related manufacturing and contributed to a shift towards services, with tourism becoming increasingly dominant as a source of growth and employment.

These features make the IPI a challenging case for seasonal adjustment. The series displays strong and persistent seasonal patterns associated with production cycles and calendar effects, a persistent level shift following 2008, and changes in volatility across different periods. Moreover, the underlying structural change in the economy complicates the interpretation of long-term movements in the trend component. Conventional methods that assume stability over the entire sample risk misattributing these permanent changes either to seasonal variation or to temporary irregular shocks, thereby distorting the adjusted series. These characteristics make the IPI an ideal candidate to illustrate the need for a seasonal adjustment procedure that explicitly accounts for structural breaks. The objective is not only to extract a seasonal component but to do so in a way that remains valid and interpretable across distinct structural regimes.

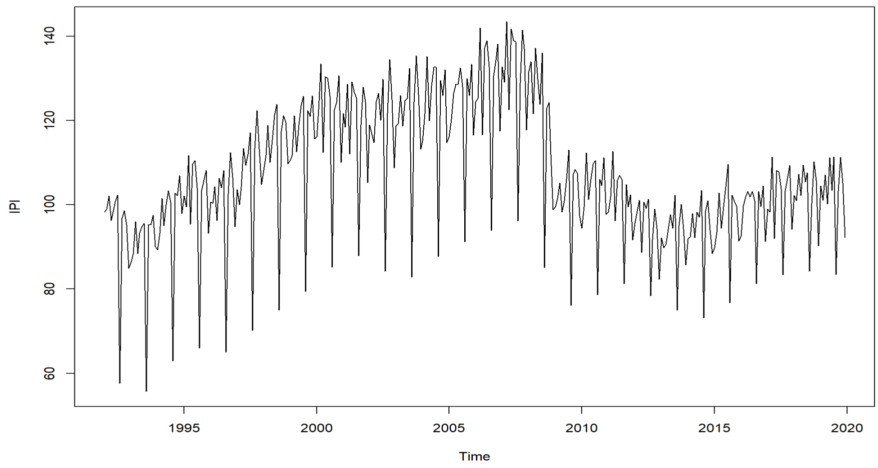

Figure 1: Spanish Industrial Production Index (IPI), raw series (1992–2019).

Our empirical strategy follows the methodological steps developed in Sections 3 and 4. First, we fit a Basic Structural Model (BSM) to the entire sample under the simplifying assumption of constant parameters. This deliberately misspecified model provides standardised innovations and smoothed disturbances that serve as diagnostic tools. Figure 2 shows the standardised innovations, together with their squared values, which reveal clear clusters of unusually high volatility around the 2008–2009 period.

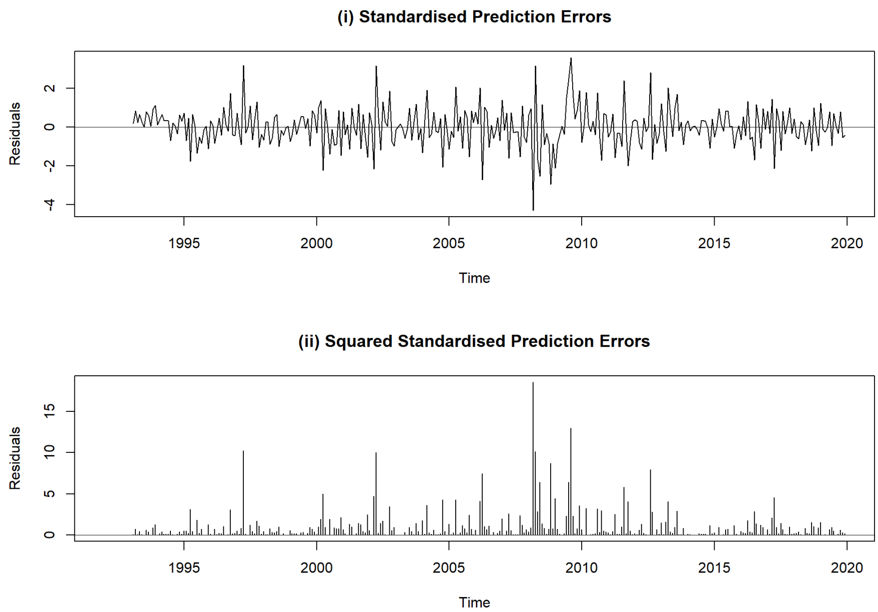

Figure 2: Regular and squared standardised innovations.

To emphasise persistent deviations, we compute rolling averages of the squared residuals with a 24-month window. The left- and right-aligned means, displayed in Figure 3, suggest the presence of a structural break window surrounding the financial crisis. We set empirical thresholds based on the 95th percentile of the rolling distribution and identify exceedances as potential indicators of instability. The lower bound a of the breakpoint region is determined from the left-aligned mean, which is more sensitive to imminent shifts, while the upper bound b is derived from the right-aligned mean, capturing the normalisation of variance post-break. The resulting interval, $[a, b] = [11/2006, 03/2010]$, is interpreted as a breakpoint region, suggesting a local structural change in variance or dynamics.

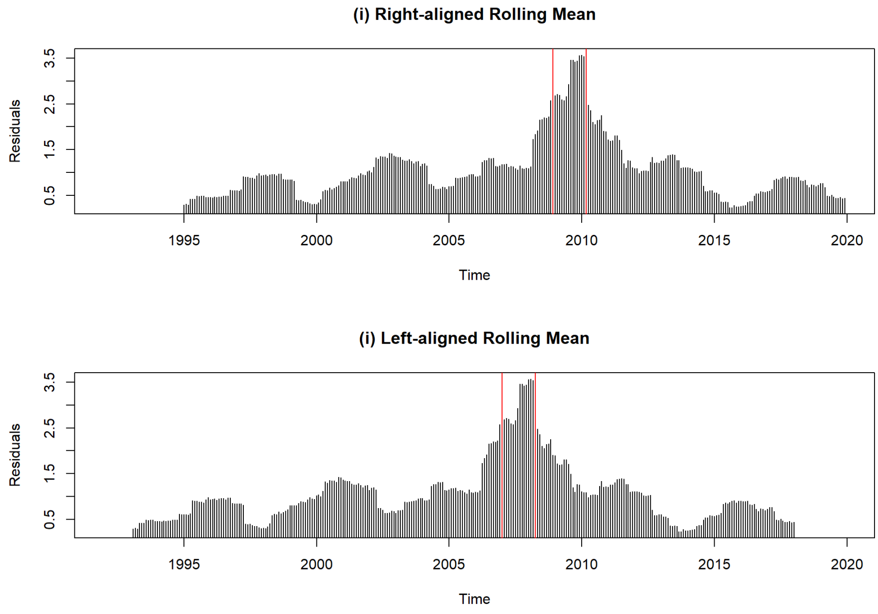

Figure 3: Left- and right-aligned rolling means of squared residuals (24-month window).

Once a structural change window has been identified, the next step is to understand the nature of the change using smoothed disturbances obtained from the Kalman smoother. We begin with a graphical analysis, Figure 4, to obtain a first clue about the type of structural disturbance. Red vertical lines mark the boundaries of the structural change window previously defined.

Figure 4: Auxiliary residuals from irregular and level components with break interval highlighted.

Next, we compare local variance estimates before, during, and after the break. Within each period, we compute the mean of squares of the smoothed disturbances. Figure 5 summarises these estimates. The results indicate that the break is predominantly associated with the level component, consistent with a structural decline in industrial output. In this application, the change is classified as abrupt in both cases.

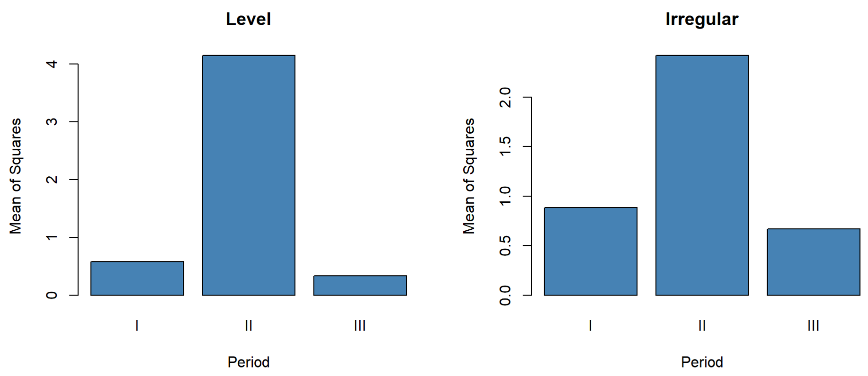

Figure 5: Local variance estimates of smoothed disturbances before, during, and after the break.

Finally, we estimate the piecewise dynamic model proposed in Section 3, which explicitly allows the variances of the state disturbances to change across regimes. Figure 6 displays the observed IPI series together with the smoothed trend, seasonal and irregular components obtained from the Kalman smoother. The decomposition clearly reflects the structural decline in the level component after the 2008 crisis, while maintaining stable seasonal patterns and a well-behaved irregular.

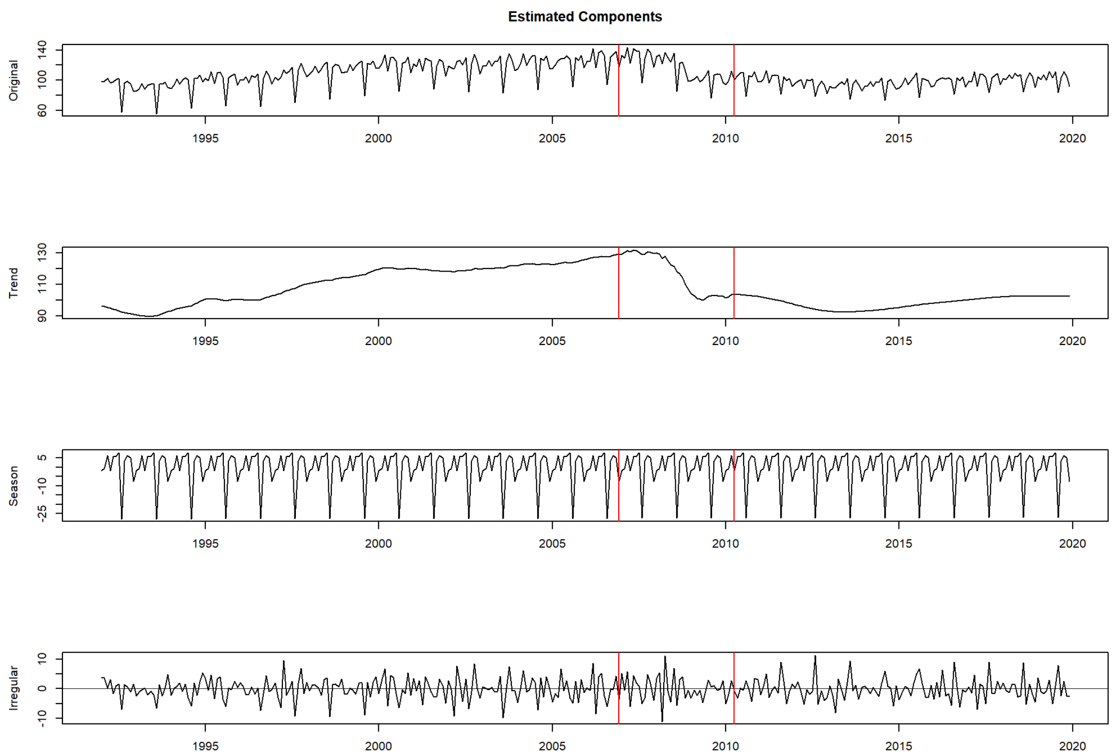

Figure 6: Observed IPI series and smoothed trend, seasonal and irregular components from the piecewise dynamic model.

6. CONCLUSIONS

This research provides a novel framework for seasonal adjustment in the presence of structural breaks, combining the flexibility of state space models with a systematic treatment of variance changes across regimes. By allowing for time-varying disturbance variances, the methodology accommodates both abrupt and gradual changes, ensuring coherent decomposition of trend, seasonal, and irregular components even in long series affected by structural instability.

A key contribution of the study is the integration of breakpoint detection and modelling into a unified framework. The detection procedure, based on rolling averages of squared innovations and smoothed disturbances, provides an intuitive and effective way to identify break regions and classify the nature of structural changes. Once identified, these changes are incorporated into the state space representation through regime-specific variance structures, preserving the statistical consistency of the decomposition and avoiding the distortions that would result from fitting separate models to unstable sub-periods.

The empirical illustration with the Spanish Industrial Production Index underscores the practical relevance of the framework for official statistics. Structural breaks driven by financial crises, long-term sectoral transformations, or institutional shifts are pervasive in socio-economic time series. If left untreated, they can bias seasonal adjustments and compromise economic interpretation. The proposed methodology provides a principled and operational solution, fully compatible with the Kalman filter and existing state space tools, thereby offering statistical offices a robust way to handle instability in the seasonal adjustment of official indicators.

ACKNOWLEDGMENTS

This work is partially funded by the INE project CIDMEFEO. It has also been funded by the Xunta de Galicia (Competitive Reference Groups, ED431C–2024/14) and by CITIC, as a center accredited for excellence within the Galician University System and a member of the CIGUS Network, and subsidized by the Regional Ministry of Education, Science, Universities and Vocational Training of the Xunta de Galicia, and co-funded through the European Regional Development Fund (ERDF), ERDF Galicia 2021–27 Operational Programme (Ref. ED431G 2023/01).

REFERENCES

Durbin, J. and Koopman, S.J. (2012). Time Series Analysis by State Space Methods. Oxford University Press.

Eurostat, 2009. ESS Guidelines on Seasonal Adjustment: 2009 edition.

Eurostat, 2015. ESS Guidelines on Seasonal Adjustment: 2015 edition.

Gómez, V. and Maravall, A., 1996. Programs tramo and seats, instruction for user (beta version: september 1996). Working papers, Banco de España.

INE, 2019. Estándar del INE para la corrección de efectos estacionales y efectos de calendario en las series coyunturales.