Les Cahiers du GERAD ISSN: 0711–2440

Linking frequentist and Bayesian change-point methods

Online Appendix

D. Ardia, A. Dufays, C. Ordás Criado

G–2023–70

December 2023

CITATION ORIGINALE / ORIGINAL CITATION

D. Ardia, A. Dufays, C. Ordás Criado (2023). "Linking Frequentist and Bayesian Change-Point Methods", *Journal of Business and Economic Statistics*, https://doi.org/10.1080/07350015.2023.2293166.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2023 – Bibliothèque et Archives Canada, 2023

Legal deposit – Bibliothèque et Archives nationales du Québec, 2023 – Library and Archives Canada, 2023

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Linking frequentist and Bayesian change-point methods

Online Appendix

David Ardia [a]

Arnaud Dufays [b]

Carlos Ordás Criado [c]

[a] *Department of Decision Sciences, HEC Montréal & GERAD, Montréal, (Qc), Canada, H3T 2A7*

[b] *Faculty of Data Science, Economics & Finance, EDHEC Business School, 59100 Roubaix, France*

[c] *Department of Economics, Laval University, Québec (Qc), Canada, G1V 0A6*

david.ardia@hec.ca
arnaud.dufays@edhec.edu

December 2023
Les Cahiers du GERAD
G–2023–70

Abstract : We show that the two-stage minimum description length (MDL) criterion widely used to estimate linear change-point (CP) models corresponds to the marginal likelihood of a Bayesian model with a specific class of prior distributions. This allows results from the frequentist and Bayesian paradigms to be bridged together. Thanks to this link, one can rely on the consistency of the number and locations of the estimated CPs and the computational efficiency of frequentist methods, and obtain a probability of observing a CP at a given time, compute model posterior probabilities, and select or combine CP methods via Bayesian posteriors. Furthermore, we adapt several CP methods to take advantage of the MDL probabilistic representation. Based on simulated data, we show that the adapted CP methods can improve structural break detection compared to state-of-the-art approaches. Finally, we empirically illustrate the usefulness of combining CP detection methods when dealing with long time series and forecasting.

Keywords : Change-point, minimum description length, model selection/combination, structural change

---

Acknowledgements: Preliminary versions of this article were presented at the 2019 annual R/Finance conference in Chicago, at the Econometric Society & Bocconi University Virtual World Congress 2020, the 35th European Economic Association Virtual Congress 2020, and the NBER-NSF Time Series Conference 2023. We benefited from helpful comments from Laurent Barras, Luc Bauwens, Keven Bluteau, Kris Boudt, Serge Darroles, Linda Mhalla, Jeroen Rombouts, and Yong Song. We are grateful to the associate editor and two anonymous reviewers for insightful comments, which helped to greatly improve the article. The usual disclaimer applies.

Data Availability Statement: Please contact the corresponding author to have access to the computer code and the datasets.

Disclosure Statement: The authors report there are no competing interests to declare.

Funding : We gratefully acknowledge the financial support of IVADO, NSERC (grant RGPIN–2022–03767), and the Fonds de recherche du Québec – Société et culture (grant #193277).

I Proofs of the Propositions

I.A Proofs of Proposition 1

The $\mathcal{NIG}$ priors are given by:

$$f(\boldsymbol{\beta}_i|\sigma_i^2,\boldsymbol{\tau}) = (2\pi)^{-\frac{K}{2}}|\sigma_i^2\underline{g}_i\underline{\mathbf{M}}_i^{-1}|^{-\frac{1}{2}}\exp\left(-\frac{1}{2\sigma_i^2}(\boldsymbol{\beta}_i-\underline{\boldsymbol{\beta}}_i)'\underline{g}_i^{-1}\underline{\mathbf{M}}_i(\boldsymbol{\beta}_i-\underline{\boldsymbol{\beta}}_i)\right), \tag{A1}$$

$$f(\sigma_i^2|\boldsymbol{\tau}) = \frac{(\frac{\underline{s}_i}{2})^{\frac{\underline{\nu}_i}{2}}}{\Gamma(\frac{\underline{\nu}_i}{2})}(\sigma_i^2)^{-\frac{\underline{\nu}_i+2}{2}}\exp\left(-\frac{\underline{s}_i}{2\sigma_i^2}\right). \tag{A2}$$

The normalizing constant of the NIG kernel is:

$$\begin{aligned} C(\underline{g}_i\underline{\mathbf{M}}_i^{-1},\frac{\underline{\nu}_i}{2},\frac{\underline{s}_i}{2}) &= \int\int(\sigma_i^2)^{-(\underline{\nu}_i+2)/2}\exp\left(-\frac{\underline{s}_i}{2\sigma_i^2}\right) \\ &\qquad (\sigma_i^2)^{-K/2}\exp\left(-\frac{1}{2\sigma_i^2}[(\boldsymbol{\beta}_i-\underline{\boldsymbol{\beta}}_i)'\underline{g}_i^{-1}\underline{\mathbf{M}}_i(\boldsymbol{\beta}_i-\underline{\boldsymbol{\beta}}_i)]\right)d\boldsymbol{\beta}_i d\sigma_i^2, \\ &= (2\pi)^{\frac{K}{2}}|\underline{g}_i\underline{\mathbf{M}}_i^{-1}|^{\frac{1}{2}}\frac{\Gamma(\frac{\underline{\nu}_i}{2})}{(\frac{\underline{s}_i}{2})^{\frac{\underline{\nu}_i}{2}}}. \end{aligned} \tag{A3}$$

The likelihood function over the segment i is given by:

$$f(\mathbf{y}_i|\boldsymbol{\tau},\boldsymbol{\beta}_i,\sigma_i^2) = (2\pi\sigma_i^2)^{-\frac{n_i}{2}}\exp\left(-\frac{1}{2\sigma_i^2}[s_i+(\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i)'\mathbf{X}_i'\mathbf{X}_i(\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i)]\right), \tag{A4}$$

where $s_i = (\mathbf{y}_i-\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i)'(\mathbf{y}_i-\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i)$. Using the above expressions, the marginal likelihood reads as follows:

$$\begin{aligned} f(\mathbf{y}_i|\boldsymbol{\tau}) = C(\underline{g}_i\underline{\mathbf{M}}_i^{-1},\frac{\underline{\nu}_i}{2},\frac{\underline{s}_i}{2})^{-1}\int\int(2\pi)^{-\frac{n_i}{2}}(\sigma_i^2)^{-(n_i+\underline{\nu}_i+K+2)/2} \\ \exp\left(-\frac{1}{2\sigma_i^2}[\underline{s}_i+s_i+\underbrace{(\boldsymbol{\beta}_i-\underline{\boldsymbol{\beta}}_i)'\underline{g}_i^{-1}\underline{\mathbf{M}}_i(\boldsymbol{\beta}_i-\underline{\boldsymbol{\beta}}_i)+(\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i)'\mathbf{X}_i'\mathbf{X}_i(\boldsymbol{\beta}_i-\widehat{\boldsymbol{\beta}}_i)}_{=F}]\right)d\boldsymbol{\beta}_i d\sigma_i^2. \end{aligned} \tag{A5}$$

Focusing on F, we can collect terms $\boldsymbol{\beta}_i$, $\underline{\boldsymbol{\beta}}_i$, and $\widehat{\boldsymbol{\beta}}_i$ as follows:

$$F = \boldsymbol{\beta}_i'\underbrace{(\underline{g}_i^{-1}\underline{\mathbf{M}}_i+\mathbf{X}_i'\mathbf{X}_i)}_{=\bar{\mathbf{M}}_i}\boldsymbol{\beta}_i - 2\boldsymbol{\beta}_i'\underbrace{(\underline{g}_i^{-1}\underline{\mathbf{M}}_i\underline{\boldsymbol{\beta}}_i+\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i)}_{=\bar{\mathbf{M}}_i\bar{\boldsymbol{\beta}}_i} + \underline{\boldsymbol{\beta}}_i'\underline{g}_i^{-1}\underline{\mathbf{M}}_i\underline{\boldsymbol{\beta}}_i + \widehat{\boldsymbol{\beta}}_i'\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i, \tag{A6}$$

where we introduce the following notation:

$$\bar{\mathbf{M}}_i = \underline{g}_i^{-1}\underline{\mathbf{M}}_i+\mathbf{X}_i'\mathbf{X}_i, \tag{A7}$$

$$\bar{\mathbf{M}}_i\bar{\boldsymbol{\beta}}_i = \underline{g}_i^{-1}\underline{\mathbf{M}}_i\underline{\boldsymbol{\beta}}_i+\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i, \tag{A8}$$

$$\bar{\boldsymbol{\beta}}_i = \bar{\mathbf{M}}_i^{-1}(\underline{g}_i^{-1}\underline{\mathbf{M}}_i\underline{\boldsymbol{\beta}}_i+\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i). \tag{A9}$$

By using (A7) to (A9), we can thus rewrite F as:

$$\begin{aligned} F &= \boldsymbol{\beta}_i'\bar{\mathbf{M}}_i\boldsymbol{\beta}_i+\underline{\boldsymbol{\beta}}_i'\underline{g}_i^{-1}\underline{\mathbf{M}}_i\underline{\boldsymbol{\beta}}_i-2\boldsymbol{\beta}_i'\bar{\mathbf{M}}_i\bar{\boldsymbol{\beta}}_i+\widehat{\boldsymbol{\beta}}_i'\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i \\ &= \underline{\boldsymbol{\beta}}_i'\underline{g}_i^{-1}\underline{\mathbf{M}}_i\underline{\boldsymbol{\beta}}_i+\widehat{\boldsymbol{\beta}}_i'\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i-\bar{\boldsymbol{\beta}}_i'\bar{\mathbf{M}}_i\bar{\boldsymbol{\beta}}_i+\boldsymbol{\beta}_i'\bar{\mathbf{M}}_i\boldsymbol{\beta}_i-2\boldsymbol{\beta}_i'\bar{\mathbf{M}}_i\bar{\boldsymbol{\beta}}_i+\bar{\boldsymbol{\beta}}_i'\bar{\mathbf{M}}_i\bar{\boldsymbol{\beta}}_i \\ &= \underline{\boldsymbol{\beta}}_i'\underline{g}_i^{-1}\underline{\mathbf{M}}_i\underline{\boldsymbol{\beta}}_i+\widehat{\boldsymbol{\beta}}_i'\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i-\bar{\boldsymbol{\beta}}_i'\bar{\mathbf{M}}_i\bar{\boldsymbol{\beta}}_i+(\boldsymbol{\beta}_i-\bar{\boldsymbol{\beta}}_i)'\bar{\mathbf{M}}_i(\boldsymbol{\beta}_i-\bar{\boldsymbol{\beta}}_i). \end{aligned} \tag{A10}$$

Using (A10) in marginal likelihood (A5) leads to:

$$
\begin{aligned}
f(\mathbf{y}_i|\boldsymbol{\tau}) &= C(\underline{g}_i\underline{\mathbf{M}}_i^{-1}, \frac{\underline{\nu}_i}{2}, \frac{\underline{s}_i}{2})^{-1} \int\int (2\pi)^{-\frac{n_i}{2}} (\sigma_i^2)^{-(n_i+\underline{\nu}_i+2)/2} \\
&\quad \exp\left(-\frac{1}{2\sigma_i^2}[\underline{s}_i + s_i + \underline{\boldsymbol{\beta}}_i'\underline{g}_i^{-1}\underline{\mathbf{M}}_i\underline{\boldsymbol{\beta}}_i + \widehat{\boldsymbol{\beta}}_i'\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i - \bar{\boldsymbol{\beta}}_i'\bar{\mathbf{M}}_i\bar{\boldsymbol{\beta}}_i]\right) \\
&\quad (\sigma_i^2)^{-\frac{K}{2}} \exp\left(-\frac{1}{2\sigma_i^2}[(\boldsymbol{\beta}_i - \bar{\boldsymbol{\beta}}_i)'\bar{\mathbf{M}}_i(\boldsymbol{\beta}_i - \bar{\boldsymbol{\beta}}_i)]\right) d\boldsymbol{\beta}_i d\sigma_i^2 \\
&= C(\underline{g}_i\underline{\mathbf{M}}_i^{-1}, \frac{\underline{\nu}_i}{2}, \frac{\underline{s}_i}{2})^{-1} (2\pi)^{-\frac{n_i-K}{2}} |\bar{\mathbf{M}}_i^{-1}|^{\frac{1}{2}} \\
&\quad \int (\sigma_i^{-2})^{(n_i+\underline{\nu}_i+2)/2} \exp\left(-\frac{1}{2\sigma_i^2}[\underline{s}_i + s_i + \underline{\boldsymbol{\beta}}_i'\underline{g}_i^{-1}\underline{\mathbf{M}}_i\underline{\boldsymbol{\beta}}_i + \widehat{\boldsymbol{\beta}}_i'\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i - \bar{\boldsymbol{\beta}}_i'\bar{\mathbf{M}}_i\bar{\boldsymbol{\beta}}_i]\right) d\sigma_i^2 \\
&= C(\underline{g}_i\underline{\mathbf{M}}_i^{-1}, \frac{\underline{\nu}_i}{2}, \frac{\underline{s}_i}{2})^{-1} (2\pi)^{-\frac{n_i-K}{2}} |\bar{\mathbf{M}}_i^{-1}|^{\frac{1}{2}} \frac{\Gamma(\frac{\bar{\nu}_i}{2})}{\frac{\bar{s}_i}{2}^{\frac{\bar{\nu}_i}{2}}} \\
&= (2\pi)^{-\frac{n_i}{2}} \left(\frac{|\bar{\mathbf{M}}_i^{-1}|}{|\underline{g}_i\underline{\mathbf{M}}_i^{-1}|}\right)^{\frac{1}{2}} \frac{\Gamma(\frac{\bar{\nu}_i}{2})}{\Gamma(\frac{\underline{\nu}_i}{2})} \frac{(\frac{\underline{s}_i}{2})^{\frac{\underline{\nu}_i}{2}}}{(\frac{\bar{s}_i}{2})^{\frac{\bar{\nu}_i}{2}}},
\end{aligned}
\tag{A11}
$$

with $\bar{\nu}_i = n_i + \underline{\nu}_i$ and $\bar{s}_i = \underline{s}_i + s_i + \underline{\boldsymbol{\beta}}_i'\underline{g}_i^{-1}\underline{\mathbf{M}}_i\underline{\boldsymbol{\beta}}_i + \widehat{\boldsymbol{\beta}}_i'\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i - \bar{\boldsymbol{\beta}}_i'\bar{\mathbf{M}}_i\bar{\boldsymbol{\beta}}_i$.

I.B Proofs of Proposition 2

If $\underline{\mathbf{M}}_i = \mathbf{X}_i'\mathbf{X}_i$ and $\underline{\boldsymbol{\beta}}_i = \widehat{\boldsymbol{\beta}}_i$, we have:

$$\bar{\mathbf{M}}_i = (1 + \underline{g}_i^{-1})\mathbf{X}_i'\mathbf{X}_i\,, \tag{A12}$$

$$\bar{\mathbf{M}}_i\bar{\boldsymbol{\beta}}_i = (1 + \underline{g}_i^{-1})\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i\,, \tag{A13}$$

$$\bar{\boldsymbol{\beta}}_i = \widehat{\boldsymbol{\beta}}_i\,, \tag{A14}$$

$$
\begin{aligned}
\bar{s}_i &= \underline{s}_i + s_i + \widehat{\boldsymbol{\beta}}_i'\underline{g}_i^{-1}\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i + \widehat{\boldsymbol{\beta}}_i'\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i - \widehat{\boldsymbol{\beta}}_i'(1 + \underline{g}_i^{-1})\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i \\
&= \underline{s}_i + s_i + \widehat{\boldsymbol{\beta}}_i'(\underline{g}_i^{-1} + 1 - 1 - \underline{g}_i^{-1})\mathbf{X}_i'\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i \\
&= \underline{s}_i + s_i\,.
\end{aligned}
\tag{A15}
$$

We can simplify the marginal likelihood of segment i as follows:

$$f(\mathbf{y}_i|\boldsymbol{\tau}) = (2\pi)^{-\frac{n_i}{2}} \left(\frac{1}{1+\underline{g}_i}\right)^{\frac{K}{2}} \frac{\Gamma(\frac{\bar{\nu}_i}{2})}{\Gamma(\frac{\underline{\nu}_i}{2})} \frac{(\frac{\underline{s}_i}{2})^{\frac{\underline{\nu}_i}{2}}}{(\frac{\bar{s}_i}{2})^{\frac{\bar{\nu}_i}{2}}}\,. \tag{A16}$$

The marginal log-likelihood of segment i reads:

$$\ln f(\mathbf{y}_i|\boldsymbol{\tau}) = -\frac{n_i}{2}\ln(2\pi) - \frac{K}{2}\ln\left(1+\underline{g}_i\right) + \ln\Gamma\left(\frac{\bar{\nu}_i}{2}\right) - \ln\Gamma\left(\frac{\underline{\nu}_i}{2}\right) + \ln\left(\frac{\underline{s}_i}{2}\right)^{\frac{\underline{\nu}_i}{2}} + \ln\left(\frac{\bar{s}_i}{2}\right)^{-\frac{\bar{\nu}_i}{2}}\,. \tag{A17}$$

Recalling Stirling's approximation of the gamma function for real $x > 0$, we have:

$$\ln\Gamma(x) = x\ln x - x - \frac{1}{2}\ln x + \frac{1}{2}\ln(2\pi) + R_N(x) + \mathcal{O}\left(x^{-(2N-1)}\right)\,, \tag{A18}$$

with $R_N(x) = \sum_{n=1}^{N-1} \frac{B_{2n}}{2n(2n-1)x^{2n-1}}$ in which B_{2n} denotes the Bernoulli numbers (*e.g.*, Nemes, 2015). Applying the Stirling's approximation as well as substituting $\underline{\nu}_i = \underline{k}_i n_i$ and $\underline{s}_i = \underline{k}_i s_i$, we find that:

$$
\begin{aligned}
\underbrace{\ln \Gamma\left(\frac{\bar{\nu}_i}{2}\right)}_{A_1} &= \frac{n_i(1+\underline{k}_i)}{2} \ln \frac{n_i(1+\underline{k}_i)}{2} - \frac{n_i(1+\underline{k}_i)}{2} - \frac{1}{2}\ln(n_i(1+\underline{k}_i)) + \frac{1}{2}\ln(4\pi) + R_N\left(\frac{n_i+\underline{\nu}_i}{2}\right) + \mathcal{O}\left((n_i+\underline{\nu}_i)^{-(2N-1)}\right), \\
\underbrace{\ln \Gamma\left(\frac{\underline{\nu}_i}{2}\right)}_{A_2} &= \frac{\underline{k}_i n_i}{2} \ln \frac{\underline{k}_i n_i}{2} - \frac{\underline{k}_i n_i}{2} - \frac{1}{2}\ln(\underline{k}_i n_i) + \frac{1}{2}\ln(4\pi) + R_N\left(\frac{\underline{k}_i n_i}{2}\right) + \mathcal{O}\left((\underline{k}_i n_i)^{-(2N-1)}\right) \\
&= \frac{n_i}{2}\ln\frac{n_i}{2} + \frac{n_i}{2}\ln(1+\underline{k}_i) + \frac{n_i\underline{k}_i}{2}\ln\frac{n_i(1+\underline{k}_i)}{2} - \frac{n_i}{2} - \frac{n_i\underline{k}_i}{2} \\
&\quad - \frac{1}{2}\ln n_i - \frac{1}{2}\ln(\underline{k}_i+1) + \frac{1}{2}\ln(4\pi) + R_N\left(\frac{n_i+\underline{\nu}_i}{2}\right) + \mathcal{O}\left((n_i+\underline{\nu}_i)^{-(2N-1)}\right), \\
\underbrace{\ln\left(\frac{\underline{s}_i}{2}\right)^{\frac{\underline{\nu}_i}{2}}}_{A_3} &= \frac{\underline{k}_i n_i}{2}\ln\frac{\underline{k}_i s_i}{2}, \\
\underbrace{\ln\left(\frac{\bar{s}_i}{2}\right)^{-\frac{\bar{\nu}_i}{2}}}_{A_4} &= -\frac{n_i(1+\underline{k}_i)}{2}\ln\frac{s_i(1+\underline{k}_i)}{2} = -\frac{n_i}{2}\ln\frac{s_i}{2} - \frac{n_i}{2}\ln(1+\underline{k}_i) - \frac{n_i\underline{k}_i}{2}\ln\frac{s_i(1+\underline{k}_i)}{2}.
\end{aligned}
\tag{A19}
$$

Using the terms in (A19), neglecting the approximation order for convenience, and recalling that $\ln f(\mathbf{y}_i|\widehat{\Theta}_{\text{MLE}}, \boldsymbol{\tau}) = -\frac{n_i}{2}\ln(2\pi\frac{s_i}{n_i}) - \frac{n_i}{2}$, we get:

$$
A_1 - A_2 + A_3 + A_4 = \ln f(\mathbf{y}_i|\widehat{\Theta}_{\text{MLE}}, \boldsymbol{\tau}) + \frac{n_i}{2}\ln(2\pi) + \frac{1}{2}\ln\frac{\underline{k}_i}{\underline{k}_i+1} + R_N\left(\frac{n_i+\underline{\nu}_i}{2}\right) - R_N\left(\frac{\underline{k}_i n_i}{2}\right),
$$

which yields:

$$
\ln f(\mathbf{y}_i|\boldsymbol{\tau}) = \ln f(\mathbf{y}_i|\widehat{\Theta}_{\text{MLE}}, \boldsymbol{\tau}) - \frac{K}{2}\ln(1+\underline{g}_i) + \frac{1}{2}\ln\frac{\underline{k}_i}{\underline{k}_i+1} + \Delta R_{N,i},
$$

where $\Delta R_{N,i} = R_N(\frac{n_i+\underline{\nu}_i}{2}) - R_N(\frac{\underline{k}_i n_i}{2})$. Summing over all segments and setting $\underline{k}_i = \frac{1}{\sqrt{n_i}}$, $\underline{g}_i = \underline{f}_i n_i - 1$ with $\underline{f}_i = \left(\frac{((m^+)^{\frac{1}{m+1}})n_i^{\frac{1}{4}}T}{(\frac{1}{\sqrt{n_i}}+1)^{\frac{1}{2}}}\right)^{\frac{2}{K}} \exp(\frac{2}{K}\Delta R_{N,i})$, we find:

$$
\begin{aligned}
\sum_{i=1}^{m+1} \ln f(\mathbf{y}_i|\boldsymbol{\tau}) &= \ln f(\mathbf{y}_{1:T}|\widehat{\Theta}_{\text{MLE}}, \boldsymbol{\tau}) - \frac{K}{2}\sum_{i=1}^{m+1}\ln n_i - \frac{K}{2}\sum_{i=1}^{m+1}\ln \underline{f}_i - \frac{1}{2}\sum_{i=1}^{m+1}\ln\frac{\frac{1}{\sqrt{n_i}}+1}{\frac{1}{\sqrt{n_i}}} + \sum_{i=1}^{m+1}\Delta R_{N,i} \\
&= \ln f(\mathbf{y}_{1:T}|\widehat{\Theta}_{\text{MLE}}, \boldsymbol{\tau}) - \frac{K+1}{2}\sum_{i=1}^{m+1}\ln n_i \\
&\quad - \frac{K}{2}\sum_{i=1}^{m+1}\ln\left(\frac{((m^+)^{\frac{1}{m+1}})T}{(\frac{1}{\sqrt{n_i}}+1)^{\frac{1}{2}}}\right)^{\frac{2}{K}} \exp\left(\frac{2}{K}\Delta R_{N,i}\right) - \frac{1}{2}\sum_{i=1}^{m+1}\ln\left(\frac{1}{\sqrt{n_i}}+1\right) + \sum_{i=1}^{m+1}\Delta R_{N,i} \\
&= \ln f(\mathbf{y}_{1:T}|\widehat{\Theta}_{\text{MLE}}, \boldsymbol{\tau}) - \frac{K+1}{2}\sum_{i=1}^{m+1}\ln n_i - \ln^+(m) - \frac{K}{2}\sum_{i=1}^{m+1}\ln(T(\exp(\Delta R_{N,i})))^{\frac{2}{K}} + \sum_{i=1}^{m+1}\ln(\exp(\Delta R_{N,i})) \\
&= \ln f(\mathbf{y}_{1:T}|\boldsymbol{\tau}, \widehat{\Theta}_{\text{MLE}}) - \ln^+(m) - (m+1)\ln T - \left(\frac{K+1}{2}\right)\sum_{i=1}^{m+1}\ln n_i \\
&= \text{MDL}(m, \boldsymbol{\tau}).
\end{aligned}
\tag{A20}
$$

The last equality holds up to an approximation order of $\sum_{i=1}^{m+1}\mathcal{O}\left((\underline{k}_i n_i)^{-(2N-1)}\right)$. When $N = 4$, it is bounded by $\mathcal{O}\left(\min_i(n_i)^{-\frac{7}{2}}\right)$ and this precision is sufficient in most applications.

I.C Proof of Remark 1

For $\underline{k}_i = \frac{1}{\sqrt{n_i}}$, we have $\underline{v}_i = \underline{k}_i n_i = \sqrt{n_i}$ and $\underline{s}_i = \underline{k}_i s_i = \frac{s_i}{\sqrt{n_i}}$. Since $\sigma_i^2|\boldsymbol{\tau} \sim IG(\frac{\underline{v}_i}{2}, \frac{\underline{s}_i}{2})$, we have:

$$\begin{aligned} E[\sigma_i^2|\boldsymbol{\tau}] = \frac{\underline{s}_i}{\underline{v}_i - 2} = \frac{\frac{s_i}{\sqrt{n_i}}}{\sqrt{n_i} - 2} = \frac{s_i}{\sqrt{n_i}(\sqrt{n_i} - 2)} = \frac{s_i}{n_i - 2\sqrt{n_i}} \quad > \sigma^2_{\text{MLE}} = \frac{s_i}{n_i}\,, \\ \text{Mode} = \frac{s_i}{\sqrt{n_i}(\sqrt{n_i} + 2)} = \frac{s_i}{n_i + 2\sqrt{n_i}} \quad < \sigma^2_{\text{MLE}} = \frac{s_i}{n_i}\,. \end{aligned} \tag{A21}$$

II Posteriors of the Bayesian estimation in Section 5

The full conditional distribution of the Gibbs sampler is given by:

$$\begin{aligned} f(\boldsymbol{\beta}_i, \sigma_i^2|y_{1:T}, \boldsymbol{\tau}) &\propto f(\mathbf{y}_i|\boldsymbol{\tau}, \boldsymbol{\beta}_i, \sigma_i^2) f(\boldsymbol{\beta}_i|\sigma_i^2, \boldsymbol{\tau}) f(\sigma_i^2|\boldsymbol{\tau}) \\ &\propto (\sigma_i^2)^{-\frac{n_i}{2}} (\sigma_i^2)^{-\frac{K}{2}} (\sigma_i^2)^{-(\frac{\underline{\nu}_i+2}{2})} \exp\left(-\frac{1}{2\sigma_i^2}(s_i + (\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i)'(1 + \underline{g}_i^{-1})(\mathbf{X}_i'\mathbf{X}_i)(\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i)) - \frac{\underline{s}_i}{2\sigma_i^2}\right) \\ &= (\sigma_i^2)^{\frac{-n_i-\underline{\nu}_i-2}{2}} (\sigma_i^2)^{-\frac{K}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i)' \frac{(1 + g_i^{-1})}{\sigma_i^2}(\mathbf{X}_i'\mathbf{X}_i)(\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i) - \frac{\underline{s}_i + s_i}{2\sigma_i^2}\right) \\ &= (\sigma_i^2)^{\frac{-(n_i+\underline{\nu}_i)}{2}-1} \exp\left(-\frac{\underline{s}_i + s_i}{2\sigma_i^2}\right) (\sigma_i^2)^{-\frac{K}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i)' \frac{(1 + g_i^{-1})}{\sigma_i^2}(\mathbf{X}_i'\mathbf{X}_i)(\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i)\right). \end{aligned} \tag{A22}$$

Direct sampling is therefore achieved as follows:

$$\begin{aligned} \sigma_i^2|y_{1:T}, \boldsymbol{\tau} &\sim \mathcal{IG}\left(\frac{n_i + \underline{\nu}_i}{2}, \frac{\underline{s}_i + s_i}{2}\right), \\ \boldsymbol{\beta}_i|y_{1:T}, \boldsymbol{\tau}, \sigma_i^2 &\sim \mathcal{N}\left(\widehat{\boldsymbol{\beta}}_i, \frac{\sigma_i^2}{(1 + \underline{g}_i^{-1})}(\mathbf{X}_i'\mathbf{X}_i)^{-1}\right), \end{aligned} \tag{A23}$$

with $\underline{\nu}_i = \sqrt{n_i}$, $\underline{s}_i = \frac{s_i}{\sqrt{n_i}}$ and $\underline{g}_i = \left(\frac{((m^+)\frac{1}{m+1})n_i^{\frac{1}{4}}T}{(\frac{1}{\sqrt{n_i}}+1)^{\frac{1}{2}}}\right)^{\frac{2}{K}} \exp(\frac{2}{K}\Delta R_{4,i})n_i - 1$ with $n_i = \tau_i - \tau_{i-1}$.

The Bayesian simulator operates as follows:

- Sample $R = 10$ initial break date vectors $\{\boldsymbol{\tau}_i\}_{i=1}^R$ from the prior distribution.
- Run an MCMC with I iterations and at each iteration, for each $j = 1, \ldots, R$, apply the D-DREAM Metropolis move:
  1. Propose a new draw of the break parameter:
  $$\mathbf{z}_j = \boldsymbol{\tau}_j + \text{round}\left[\gamma(\delta, m)\left(\sum_{g=1}^{\delta} \boldsymbol{\tau}_{r_1(g)} - \sum_{h=1}^{\delta} \boldsymbol{\tau}_{r_2(h)}\right) + \xi\right],$$
  with $\forall g, h = 1, 2, \ldots, \delta$ and $j \neq r_1(g)$, $r_2(h)$; $r_1(.)$ and $r_2(.)$ denote random integers uniformly distributed over support $[1, R]$. The round$[\cdot]$ operator picks the nearest integer and $\xi \sim \mathcal{N}(0, (0.0001)I)$. We set $\gamma(\delta, m) = \frac{2.38}{\sqrt{2\delta m}}$ and $\delta \sim \mathcal{U}(1, 3)$.
  2. Accept the proposal $\mathbf{z}_j$ with probability:
  $$\min\left\{\frac{f(\mathbf{z}_j|\mathbf{y}_{1:T})}{f(\boldsymbol{\tau}_j|\mathbf{y}_{1:T})}, 1\right\}.$$

In practice, we set $I = 1000$ and start collecting the draws after $\frac{I}{2}$ iterations.

III Additional empirical analyses

This section provides additional analyses for the two empirical illustrations presented in Section 6 of the paper.

III.A Combination of local CP methods in Fama-French factor models

To show that our CP method does not only capture time-varying volatility, we apply our CP approach to the residuals of a linear five-factor model with GARCH(1,1) errors estimated by maximum likelihood. Results are displayed in Figure A.1. We see that CPs are still observed when fitting the model on the homoscedastic residuals.

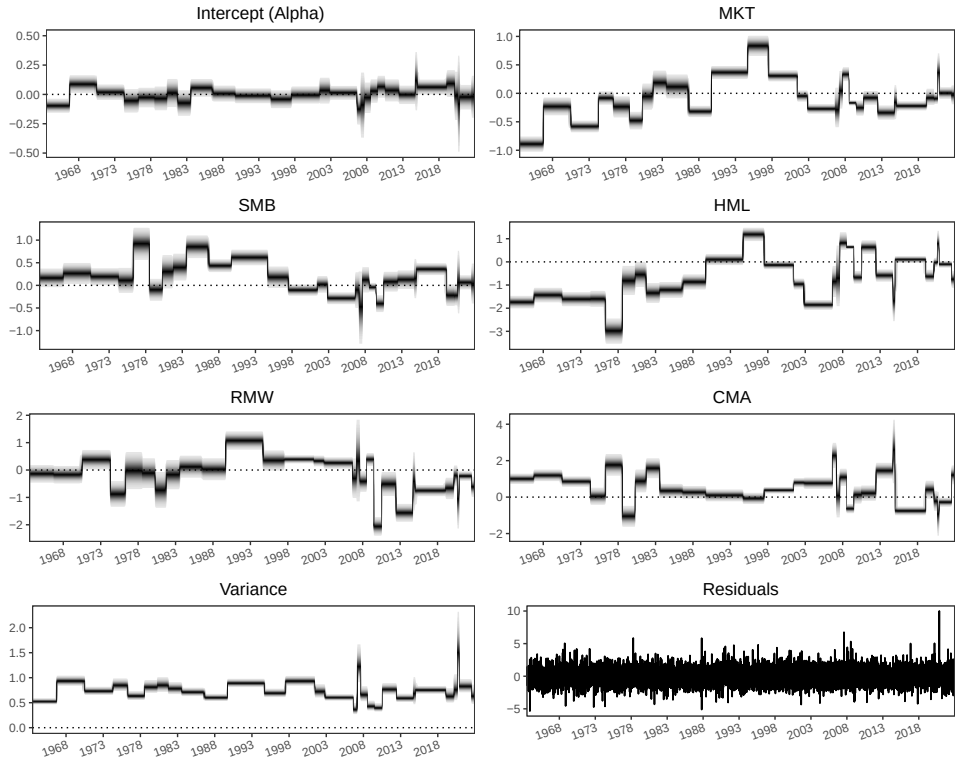

Figure A.1: Mixture of marginal posterior distributions—residuals from GARCH filtering. The plots show the mixture of marginal posteriors of the Fama-French five-factor model parameters applied to the residuals of the Fama-French five-factor model with GARCH(1,1) errors estimated by maximum likelihood on the initial time series. The bottom-right plot displays the series of residuals.

We now show empirically that our prior setting does not noticeably impact the posterior distribution of the parameter in our Fama-French application. We consider a CP model in which the CP locations are set to the estimated values found by the global approach (*i.e.*, the CPs estimated by the best model reported in Table 6 of the paper). For this model, we specify the priors for the mean and variance parameters as:

$$\begin{aligned} \boldsymbol{\beta}_i &\sim \mathcal{N}\left(\mathbf{0}, 100 I_K\right) , \\ f(\sigma_i^2) &\propto \frac{1}{\sigma_i^2} , \end{aligned} \tag{A24}$$

where I_K denotes the K-dimensional identity matrix. We then apply 10,000 MCMC iterations for sampling the mean and the variance parameters of each segment from their full conditional distributions (since the posterior distribution is not a Normal-Inverse Gamma distribution).

In Figures A.2–A.3 below, we display the comparison between the posterior distribution of the model parameters obtained with the NIG prior (left plots) and the proper but non-informative prior (right plots) given in Equation (A24) above. We display the median (solid line) and the 90th credible interval (dashed lines). We see that results obtained from each prior are almost undistinguishable.

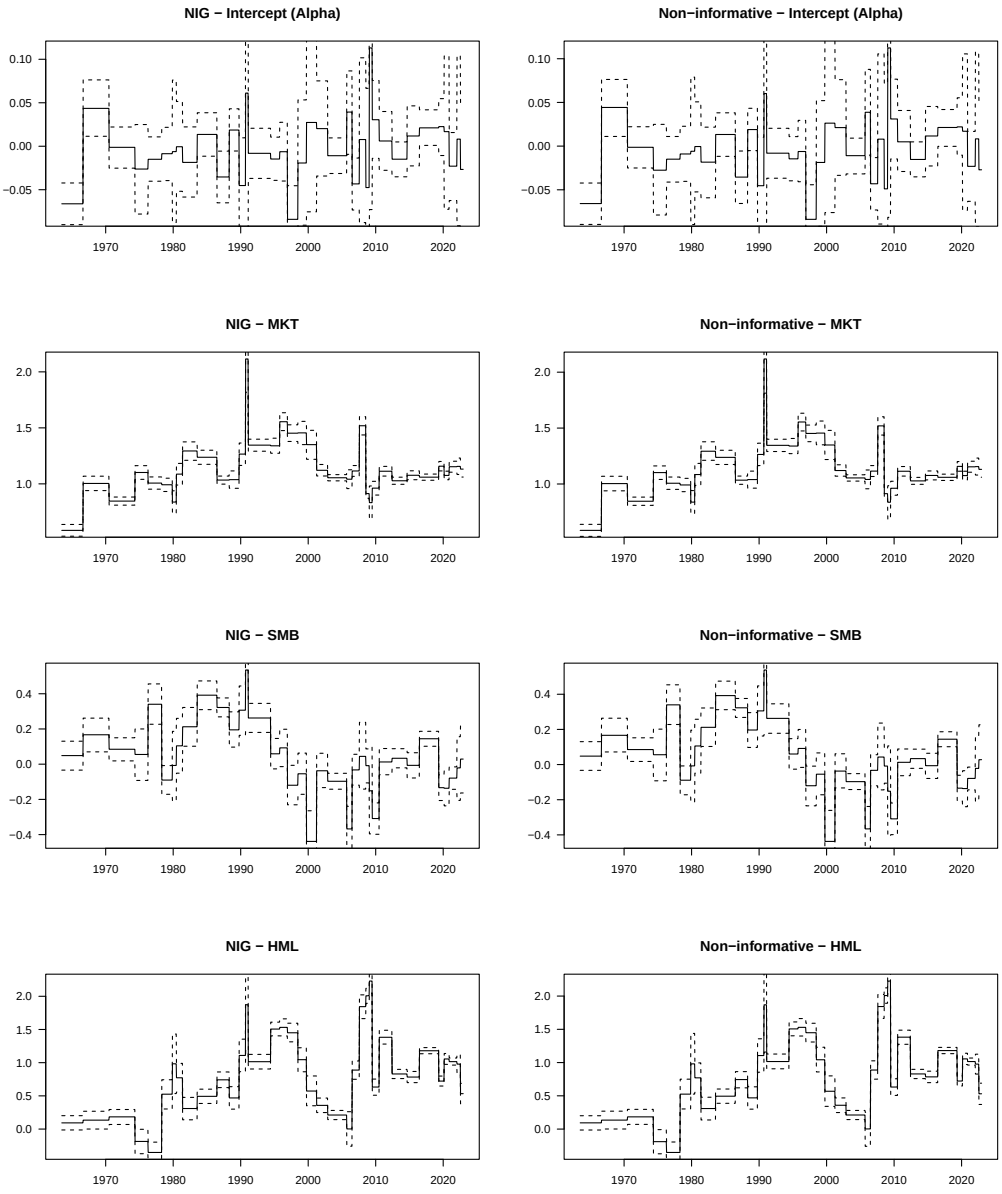

Figure A.2: Impact of priors on marginal posterior distributions—first set of parameters. The plots show the marginal posterior distributions (conditional on the CPs) of the various model parameters for the NIG prior (left) and the proper but non-informative prior (right). The solid line depicts the median of the posterior and the dashed lines depict the 95% credible interval.

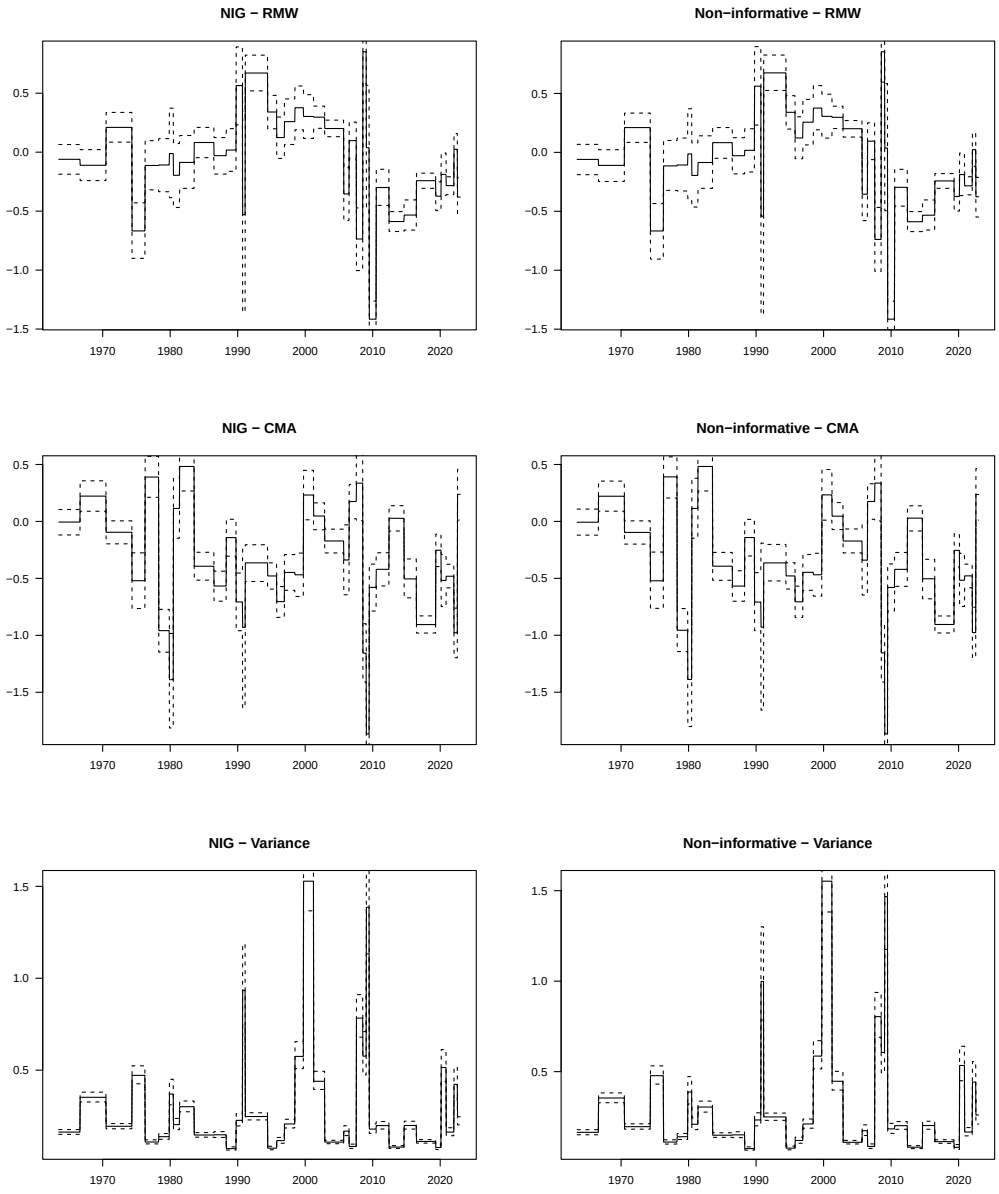

Figure A.3: Impact of priors on marginal posterior distributions—second set of parameters. The plots show the marginal posterior distributions (conditional on the CPs) of the various model parameters for the NIG prior (left) and the proper but non-informative prior (right). The solid line depicts the median of the posterior and the dashed lines depict the 95% credible interval.

III.B Forecasting the U.S. consumer price index

To assess the impact of longer-lag CP-AR specifications on the forecasting performance, we consider two additional models: the AR(3) and AR(1,2,3,12) specifications, with and without breaks. Results are reported in Table A.1. The AR(3) model outperforms its AR counterparts with no breaks. Nevertheless, the CP-AR(2) model remains the best prediction model in terms of both RMSFE and MAFE across forecasting horizons. Notably, transitioning from a local to a global approach or incorporating a future break improves forecasting performance (with exceptions for the AR(3) - Local - Future Breaks for $h = 6$). Note that the model confidence used below changes when less or more models are competing. Therefore, the gray entries in Table 7 of the paper slightly differ from Table A.1.

Table A.1: Forecasting results—additional AR specifications.

	RMSFE for horizon $h =$				MAFE for horizon $h =$			
Model	1	3	6	12	1	3	6	12
AR(1)	3.00	3.60	3.77	3.83	2.16	2.56	2.67	2.72
CP-AR(1) - Local	2.96	3.28	3.41	3.59	2.08	2.32	2.42	2.59
CP-AR(1) - Local - Future Break	2.95	3.28	3.40	3.57	2.08	2.32	2.41	2.57
CP-AR(1) - Global	2.95	3.24	3.40	3.57	2.07	2.30	2.40	2.54
CP-AR(1) - Global - Future Break	2.95	3.24	3.39	3.54	2.07	2.29	2.39	2.52
AR(2)	2.98	3.45	3.65	3.81	2.12	2.45	2.57	2.70
CP-AR(2) - Local	2.91	3.29	3.41	3.58	2.05	2.32	2.41	2.58
CP-AR(2) - Local - Future Break	2.91	3.28	3.40	3.56	2.06	2.32	2.41	2.56
CP-AR(2) - Global	2.94	3.29	3.40	3.58	2.07	2.31	2.40	2.57
CP-AR(2) - Global - Future Break	2.94	3.28	3.38	3.53	2.07	2.31	2.38	2.53
AR(3)	2.93	3.35	3.50	3.74	2.07	2.36	2.46	2.65
CP-AR(3) - Local	2.92	3.33	3.43	3.66	2.06	2.35	2.40	2.59
CP-AR(3) - Local - Future Break	2.92	3.33	3.42	3.64	2.06	2.34	2.40	2.58
CP-AR(3) - Global	2.92	3.32	3.42	3.65	2.06	2.34	2.40	2.58
CP-AR(3) - Global - Future Break	2.92	3.31	3.41	3.62	2.06	2.34	2.39	2.56
AR(1,2,3,12)	2.96	3.47	3.73	4.14	2.10	2.47	2.65	2.96
CP-AR(1,2,3,12) - Local	2.94	3.41	3.60	3.88	2.08	2.41	2.53	2.75
CP-AR(1,2,3,12) - Local - Future Break	2.94	3.41	3.59	3.85	2.08	2.40	2.52	2.73
CP-AR(1,2,3,12) - Global	2.94	3.40	3.56	3.84	2.08	2.40	2.49	2.71
CP-AR(1,2,3,12) - Global - Future Break	2.93	3.39	3.55	3.81	2.08	2.40	2.48	2.69

NOTE: This table reports the root mean square forecast error (RMSFE) and mean absolute forecasting error (MAFE) for the various model specifications and forecasting horizons. We highlight in gray the specifications belonging to the set of superior models at the 95% confidence level for a given loss function and forecasting horizon (Hansen et al., 2011). We use the function `mcsTest` implemented in the R package **rugarch** (Ghalanos, 2022). The forecasting exercise is conducted over 817 out-of-sample observations.

Table A.2 reports the CPs corresponding to the five largest MDL marginal log-likelihoods for the eleven local CP methods (Panel A) and the global method (Panel B) obtained for the CP-AR(2) specification over the full sample (in-sample).

Table A.2: Best local CP and GMDL models—AR(2) model full-sample.

	Panel A: Local CP Approaches				
	Ranking				
	Best	2nd	3rd	4th	5th
Number of CPs	9	9	11	9	10
MDL MLL	-2,283.29	-2,285.83	-2,286.42	-2,284.49	-2,284.62
Posterior Probability	56.68%	20.98%	6.31%	5.54%	5.17%
	Panel B: GMDL Approach				
	Ranking				
	Best	2nd	3rd	4th	5th
Number of CPs	10	9	11	10	12
MDL MLL	-2,276.84	-2,276.91	-2,277.60	-2,279.21	-2,279.80
Posterior Probability	38.90%	36.59%	18.33%	3.67%	2.03%

NOTE: This table reports the five best local CP and GMDL models with respect to the MDL marginal log-likelihood. Panel A reports the top five models (number of CPs) obtained among the 11 local CP methods for the AR(2) specification. Panel B reports the top five CPs for the GMDL approach. The GMDL approach is estimated with a maximum of 80 regimes.

Finally, the existing literature has documented a lot the time-varying properties of U.S. inflation, see, *e.g.*, Groen et al. (2013) and the literature therein. In Figure A.4, we display the CPI series (top left) and the mixture of the marginal posterior distributions for the unconditional mean (top right), persistence (sum of AR parameters, bottom left), and variance (bottom right) obtained over the entire time period for the global CP-AR(2) model. We notice many CPs in the (functions of the)

AR parameters. As Groen et al. (2013, section 4.3) emphasize : "The persistence of U.S. inflation increased during the 1970s, reaching peaks around 1974-1975 and around 1980, and subsided after 1982-1983 (...)." This is what we observe as well in our monthly U.S. CPI inflation series in the bottom-left panel. In our case, persistence also increased during the 2007-2008 financial crisis and further increased at the end of our sample. These authors also report a "downward shift in the U.S. inflation variance from the late 1980s and early 1990s (...) that also happened during the 1970s (...)." These features are clearly apparent in the bottom-right panel. In addition, we capture a large drop in inflation variance after WWII, an upward spike in inflation variance during the first oil crisis followed by a drop between 1975-1976, and upward shifts during the 2007-2008 financial crisis and at the end of the sample.

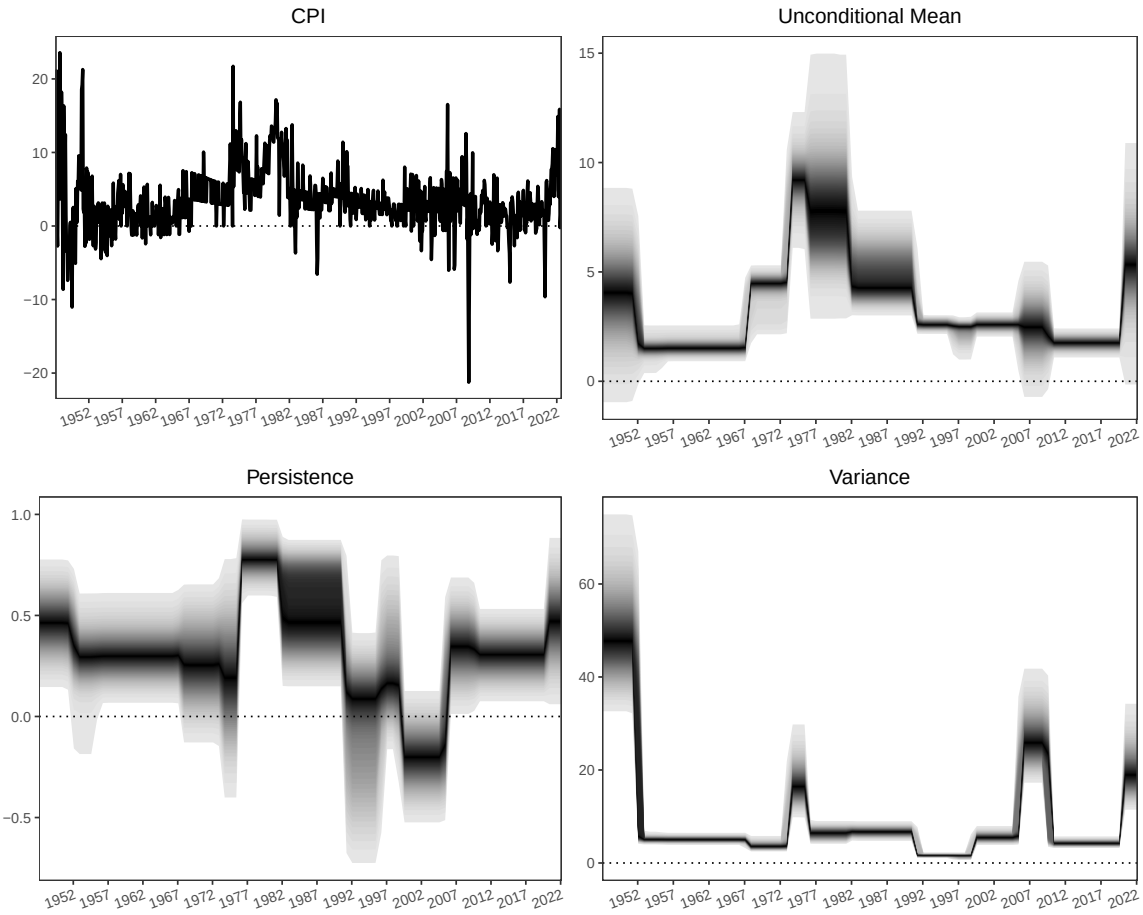

Figure A.4: CPI series and mixture of marginal posterior distributions of (functions of) the global CP-AR(2) model parameters. The plots show the CPI series (top left) and the mixture of the marginal posterior distributions for the unconditional mean (top right), persistence (sum of AR parameters, bottom left), and variance (bottom right) obtained over the entire time period for the global CP-AR(2) model. We can notice many CPs in the (functions of the) parameters. These posteriors are computed with the full Bayesian setup of Section 5. The mixing weights are the posterior probabilities of the eleven CP local methods.

References

Ghalanos, A., 2022. rugarch: Univariate GARCH models. R package version 1.4-9.

Groen, J.J.J., Paap, R., Ravazzolo, F., 2013. Real-time inflation forecasting in a changing world. Journal of Business & Economic Statistics 31, 29–44.

Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. Econometrica 79, 453–497.

Nemes, G., 2015. Error bounds and exponential improvements for the asymptotic expansions of the gamma function and its reciprocal. Proceedings of the Royal Society of Edinburgh Section A: Mathematics 145, 571–596.