CONTRIBUTI DI RICERCA CRENOS

INFLUENCE OF LOCAL AND GLOBAL ECONOMIC POLICY UNCERTAINTY ON THE VOLATILITY OF US STATE-LEVEL EQUITY RETURNS: EVIDENCE FROM A GARCH-MIDAS APPROACH WITH SHRINKAGE AND CLUSTER ANALYSIS

Vincenzo Candila
Oguzhan Cepni
Giampiero M. Gallo
Rangan Gupta

WORKING PAPERS

2024/14

arkadia

CENTRO RICERCHE ECONOMICHE NORD SUD
(CRENoS)
UNIVERSITÀ DI CAGLIARI
UNIVERSITÀ DI SASSARI

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.unica.it
crenos@unica.it

CRENoS – CAGLIARI
VIA SAN GIORGIO 12, I-09124 CAGLIARI, ITALIA
TEL. +39-070-6756397

CRENoS - SASSARI
VIA MURONI 25, I-07100 SASSARI, ITALIA
TEL. +39-079-213511

Title: INFLUENCE OF LOCAL AND GLOBAL ECONOMIC POLICY UNCERTAINTY ON THE VOLATILITY OF US STATE-LEVEL EQUITY RETURNS: EVIDENCE FROM A GARCH-MIDAS APPROACH WITH SHRINKAGE AND CLUSTER ANALYSIS

Prima Edizione: Settembre 2024
ISBN: 978 88 68515 416

Viale Bonaria 98 - 09125 Cagliari
Tel. 070/6848663 - info@arkadiaeditore.it
www.arkadiaeditore.it

Influence of Local and Global Economic Policy Uncertainty on the volatility of US state-level equity returns: Evidence from a GARCH-MIDAS approach with Shrinkage and Cluster Analysis

Vincenzo Candila
Department of Economics and Statistics, University of Salerno, Italy
vcandila@unisa.it
Oguzhan Cepni
Ostim Technical University, Ankara, Turkiye
Department of Economics, Copenhagen Business School, Denmark
oce.eco@cbs.dk
Giampiero M. Gallo
Italian Court of Audits (Corte dei conti), NYU in Florence, Italy, and CRENoS
giampiero.gallo@nyu.edu
Rangan Gupta
Department of Economics, University of Pretoria, South Africa
rangan.gupta@up.ac.za

September 25, 2024

Abstract
This paper examines the influence of local (state-specific) and global Economic Policy Uncertainty (EPU) on the volatility of US state-level equity returns. We employ a GARCH-MIDAS approach that incorporates multiple EPU indices as low-frequency predictors of daily stock return volatility. To address the challenge of selecting the most relevant EPU indices, we utilize an Elastic Net (EN) shrinkage method to combine forecasts from different models. Our results reveal that the combined model, which leverages information from both local and global EPU indices, generally outperforms single specifications. Further, a cluster analysis based on the volatility forecasts uncovers distinct geographical patterns, suggesting that state-level volatility is influenced by both state-specific and nationwide policy uncertainties. These findings highlight the importance of considering both local and global economic policy uncertainty in understanding and predicting the volatility dynamics at the regional level.

Keywords: GARCH-MIDAS, Economic Policy Uncertainty, Elastic Net, Forecast Combination, Cluster Analysis
Jel Codes: C32, C53, G10, G17, D80

1 Introduction

The present value model of asset prices (Shiller, 1981a,b) suggests that asset market volatility depends on the variability of cash flows and the discount factor. At the same time, general equilibrium models developed by Pástor and Veronesi (2012, 2013) shed light on the role played by uncertainty about government policy; they point out that policy changes raise the volatility of the stochastic discount factor, which, in turn, causes risk premia to go up and stock returns to become more volatile. By the same token, increases in policy-related uncertainties can also lead to a reduction in volatility due to lower trading volume in the equity market, as per the Mixture of Distribution Hypothesis (MDH, Clark, 1973) or the Sequential Information Arrival Hypothesis (EN, Copeland, 1976), since investors are then likely to move into "safe haven" assets (Balcilar et al., 2016; Raza et al., 2018). Hence, uncertainty surrounding government policy decisions can be associated with either an increase or a decrease in stock market volatility.

Given these transmission mechanisms, some researchers (see, for example, Liu and Zhang (2015); Liu et al. (2017); Gong et al. (2022); Li et al. (2023); Salisu et al. (2023, 2024b)) have resorted to the index of Economic Policy Uncertainty (EPU) constructed by Baker et al. (2016) from newspaper articles, to successfully forecast aggregate stock market volatility of the United States (US) by employing primarily variants of the generalized autoregressive conditional heteroskedasticity (GARCH) model (Bollerslev, 1986). Unlike the studies that use aggregated stocks and national-level EPU, we focus, for the first time, on individual state-level analysis using a new dataset developed by Baker et al. (2022) on state-level EPU. The state-level equity returns are derived from the sub-aggregation of stocks returns of firms within each of the 50 US states being considered, based on the location of their headquarters. The rationale for taking such a regional perspective is derived from the premise that core business activities of firms often occur close to their headquarters (Pirinsky and Wang, 2006; Chaney et al., 2012) and, hence, equity prices should contain a non-negligible regional (own- and neighbouring-state) component, so much so that investors' portfolios over-represent local firms (Coval and Moskowitz, 1999, 2001; Korniotis and Kumar, 2013). Nationally aggregated data tends to overlook the heterogeneous nature of the states, potentially failing to capture the true dynamics within specific groups of states. Analyzing data at the state-level thus highlights the unique characteristics and variations among individual states, revealing their specific dynamics in the examined stock returns volatility–EPU nexus. Obviously then, the forecasting exercise we undertake in this research turns out to be relevant for investors, in that it has valuable implications for portfolio selection, derivative pricing, risk management, and also for policy-making (Poon and Granger, 2003; Rapach et al., 2008).

The range of return variability affecting asset prices is time-varying and tends to be persistent through time. This so-called phenomenon of volatility clustering (Engle, 1982) is at the basis of a massive literature in financial econometrics, comprising both measurement and modeling of volatility. With relatively recent developments in exploiting the ultra-high frequency data on market prices, measurement is accomplished by building one of the many versions of realized volatility (Andersen and Benzoni, 2009), an end-of-day observed variable the dynamics of which can be modeled with either additive (e.g. Corsi, 2009) or multiplicative errors (Engle, 2002; Cipollini et al., 2021). In the absence of data on realized volatility, the GARCH literature still receives substantial empirical attention, as it simply uses the time series of close-to-close daily returns. The phases of measurement and modeling are accomplished at once: the signal in the conditional mean of the returns is treated as negligible, and the object of interest is the dynamics of the conditional variance, possibly accommodating an asymmetric reaction of volatility to negative returns (as in the GJR model of Glosten et al., 1993, for instance). While an autoregressive dependence on past squared returns is common to various GARCH approaches, there has been an increasing awareness about the presence of a slow-moving component following the paper by Engle and Rangel (2008) who recognize that there is a time-varying average level of volatility that moves at a low frequency around which the short-run dynamics evolve. Amado et al. (2019) provide a survey of models in the GARCH family addressing this empirical regularity: next to the spline-GARCH by Engle and Rangel (2008), the smooth transition GARCH of Amado and Teräsvirta (2008) and, most importantly for what is covered in this paper, the GARCH-mixed data sampling (MIDAS) by Engle et al. (2013). This model, in particular, has the feature of filtering out information observed at a lower frequency to provide a contribution to the volatility dynamics: such information may be economic-based (e.g. the industrial production to proxy phases of the business cycle, as in Conrad and Loch, 2015), financial-based (e.g. volatility observed at monthly level, as in Fang et al., 2020), or sentiment-based as in the monthly EPUs considered in what follows. A generalization to the original GARCH-MIDAS is given by Amendola et al. (2019), with a detailed review of this approach provided in Amendola et al. (2021), Segnon et al. (2024), and Salisu et al. (2024a).

In our econometric analyses, we adopt the same GARCH-MIDAS specification and we utilize alternative variants of the information from different low-frequency local and global EPU indices in the form of eleven monthly EPU-based predictors, next to the GJR model. To simplify matters, while adaptively identify the most informative EPU indices, we adopt an Elastic Net (EN, Zou and Hastie, 2005) shrinkage combination strategy of the volatility forecasts. To the best of our knowledge, this is the first time that a EN-based combination approach is used within the class of the GARCH-MIDAS models. Interestingly, we find that the shrinkage combination approach typically yields volatility forecasts that are superior or at least equivalent to those of the GJR and all the other GARCH-MIDAS models here used, as determined by the Set of Superior Models (SSM) of the Model Confidence Set (MCS, Hansen et al., 2011). These results are robust to changes in different parameters like the loss function employed in the MCS or the frequency at which the combined predictor updates.

To gain further insights on the meaning of the results, we perform a novel cluster analysis to identify homogeneous regions sharing similar volatility forecast features. In particular, following Luo et al. (2023) and references therein, we use the hierarchical clustering with Dynamic Time Warping (DTW, Sakoe and Chiba, 1978) as the distance metric and Ward's method for linkage (Ward Jr, 1963; Murtagh and Legendre, 2014). Commonly used validity indices are employed to check the goodness of the generated clusters, as in Nanda et al. (2010) and Chaudhuri and Ghosh (2015), among others. Notably, the clusters obtained from the EN-based combination approach have better validity indices with respect to the clusters generated by the simpler GJR model without the low-frequency component.

The rest of the paper is as follows. Section 2 presents the methodology employed, with the GARCH-MIDAS model formulation illustrated in 2.1 and the EN-based combination strategy in 2.2. Section 3 is devoted to the empirical analysis, and Section 4 concludes.

2 Methodology

2.1 GARCH-MIDAS model

The framework of this work relies on the general autoregressive heteroskedastic formulation with a zero conditional mean for the daily log-returns $r_{i,t}$ that is:

$$r_{i,t}|\mathscr{F}_{i-1,t} = \sigma_{i,t}\eta_{i,t}, \tag{1}$$

where i represents the day of the low-frequency period t, with $i = 1, \cdots, N_t$, and $t = 1, \cdots, T$, N_t is the number of daily observations within the low-frequency period t and T is the total number of low-frequency periods, with N denoting the total number of days, $\mathscr{F}_{i-1,t}$ is the information set up to day $i-1$, $\sigma_{i,t}$ is the conditional standard deviation of $r_{i,t}$, and $\eta_{i,t}$ is the error term, with $\eta_{i,t} \overset{i.i.d}{\sim} \mathscr{N}(0,1)$. In order to accommodate the double multiplicative component in Eq. (1), we use the GARCH-MIDAS model, for which the conditional standard deviation $\sigma_{i,t}$ decomposes into a long- and short-run components, namely

$\sigma_{i,t} = \sqrt{\tau_t g_{i,t}}$. In particular, the short-run component follows a unit-mean GJR-GARCH(1,1) process, that is:

$$g_{i,t} = (1-\alpha-\beta-\gamma/2) + \frac{\left(\alpha + \gamma \mathbb{1}_{(r_{i-1,t}<0)}\right) r_{i-1,t}^2}{\tau_t} + \beta g_{i-1,t}, \tag{2}$$

with $\mathbb{1}_{(\cdot)}$ denoting an indicator function which is equal one when the argument is true. The long-run τ_t, varying each period t, is specified as

$$\tau_t = \exp\left\{ m + \theta \sum_{k=1}^{K} \delta(\omega) MV_{t-k} \right\}, \tag{3}$$

where MV_t is the additional exogenous variable, whose K lagged realizations are weighted according to the weighting function $\delta(\omega)$. In this work, the Beta function will be used as a weighting function, defined as follows:

$$\delta_k(\omega) = \frac{(k/K)^{\omega_1-1}(1-k/K)^{\omega_2-1}}{\sum_{j=1}^{K}(j/K)^{\omega_1-1}(1-j/K)^{\omega_2-1}} \tag{4}$$

subject to the constraint $\omega_1 = 1$, to allow for a monotonically decreasing weighting scheme, where the most recent observations are given more importance. Other weighting functions are discussed in Ghysels and Qian (2019).

The parameter space of the GARCH-MIDAS model presented above is $\Theta = \{\alpha, \beta, \gamma, m, \theta, \omega_2\}$. Given the normality assumption for the error term in (1), the following log-likelihood is to be maximized:

$$\mathscr{L}(\Theta) = -\frac{1}{2} \sum_{t=1}^{T} \sum_{i=1}^{N_t} \left[\log(2\pi) + \log(g_{i,t}\tau_t) + \frac{r_{i,t}^2}{g_{i,t}\tau_t} \right]. \tag{5}$$

In addition to the GARCH-MIDAS, we also use the GJR model. In coherence with the double index formulation used in Eq. (1), the conditional variance of $r_{i,t}$, $\sigma_{i,t}^2$, is modelled as:

$$\sigma_{i,t}^2 = \omega_0 + \left(\alpha + \gamma \mathbb{1}_{(r_{i-1,t}<0)}\right) r_{i-1,t}^2 + \beta \sigma_{i-1,t}^2. \tag{6}$$

2.2 Shrinkage combination strategy

The literature on forecast combinations has its origins in the paper by Bates and Granger (1969) (see, for a review, Timmermann, 2006; Wang et al., 2023b, among others). Combination of volatility forecasts has its own peculiarities addressed in the key contributions by Amendola and Storti (2008) in the univariate and Amendola and Storti (2015) in the multivariate context. Within the GARCH-MIDAS approach, Asgharian et al. (2013) combined different sources of information, via principal component analysis to select the most relevant macroeconomic variables for the S&P 500 long-run and overall volatilities. For the same index volatility, Fang et al. (2020) resorted to the Adaptive Least Absolute Shrinkage and Selection Operator (LASSO, Zou, 2006) to combine twenty low-frequency macroeconomic and financial variables via the Adaptive-LASSO. Shrinking methods to combine forecasts from different volatility models are quite common in general: for instance, Zhang et al. (2019) used the LASSO (Tibshirani, 1996) and Elastic Net (EN, Zou and Hastie, 2005) methods in the context of Heterogeneous AutoRegressive (HAR, Corsi, 2009) model to forecast the oil price volatility. Recently, Wang et al. (2023a) used LASSO and EN to identify the most important Bitcoin volatility drivers among several macroeconomic and technical factors.

In order to adaptively weight the best performing models using (as in GARCH-MIDAS) the low-frequency additional covariates, or not using such additional variables, we resort to a combination strategy across J models. The adopted combination strategy is based on the EN shrinkage method, which uses both the L_1 and L_2 penalty functions. The L_1 penalty function refers to the Least Absolute Shrinkage and Selection (LASSO, Tibshirani, 1996) method:

$$\text{LASSO}: \quad \arg\min_{\beta} \left(\frac{1}{2N} \sum_{t=1}^{T} \sum_{i=1}^{N_t} \left(\sigma_{i,t}^2 - \beta_0 - \sum_{j=1}^{J} \beta_j \sigma_{i,t,j}^2 \right)^2 + \lambda \sum_{j=1}^{J} |\beta_j| \right), \tag{7}$$

j identifies the j-th model, with $j = 1, \cdots, J$; $\sigma_{i,t,j}^2$ the conditional variance from model j, and λ is the nonnegative regularization parameter controlling the strength of the L_1 penalization function (the second addendum in (7)). In particular, higher λs imply stronger regularization, that is, more coefficients are shrunk towards zero or even set exactly to zero.

The L_2 penalty function refers the Ridge regression (Hoerl and Kennard, 1970):

$$\text{RIDGE}: \quad \arg\min_{\beta} \left(\frac{1}{2N} \sum_{t=1}^{T} \sum_{i=1}^{N_t} \left(\sigma_{i,t}^2 - \beta_0 - \sum_{j=1}^{J} \beta_j \sigma_{i,t,j}^2 \right)^2 + \lambda \sum_{j=1}^{J} \beta_j^2 \right). \tag{8}$$

Contrary to the L_1 penalty function in (7) which penalizes the sum of coefficients in absolute values, the L_2 penalty function in (8) penalizes the sum of the squares of the coefficients (namely, the second addendum in (8)).

Specifically designed for dealing with high correlated predictors, the EN is then as follows:

$$\text{EN}: \quad \arg\min_{\beta} \left(\frac{1}{2N} \sum_{t=1}^{T} \sum_{i=1}^{N_t} \left(\sigma_{i,t}^2 - \beta_0 - \sum_{j=1}^{J} \beta_j \sigma_{i,t,j}^2 \right)^2 + \lambda \left(\rho \sum_{j=1}^{J} |\beta_j| + (1-\rho) \sum_{j=1}^{J} \beta_j^2 \right) \right), \tag{9}$$

where ρ is the mixing parameter controlling the balance between L_1 and L_2 penalties. In this work, we fix $\rho = 0.5$.

Our idea is to use the EN as a combination strategy to adaptively identify the superior models, among several GARCH-MIDAS models and the GJR specification, taking benefit from different low-frequency variables whose informative content towards the dependent variable of interest could be time-varying. For instance, if the informative power of a subset of low-frequency variables included within the GARCH-MIDAS models is low for a given period of time, the EN will shrink to zero the importance of these models, for that period.

Following the setup in Amendola et al. (2020), we define the *training* (or in-sample) period as the sample used to find the optimal $\widehat{\boldsymbol{\beta}}$ via (9). For the *testing* (or out-of-sample) period, starting on the day following the last available day of the *training* sample, N_T, T, and ending after H days (without re-estimating), the combined one-step ahead conditional variance forecasts, named EN-Comb, will be calculated as:

$$\widehat{\sigma}_{N_T+h,T}^{2,(Comb)} = \boldsymbol{X}'_{N_T+h,T} \widehat{\boldsymbol{\beta}}, \quad h = 1, \ldots, H \tag{10}$$

where, for each $h = 1, \ldots, H$, $\boldsymbol{X}_{N_T+h,T}$ represents the $J \times 1$ vector of one-step ahead conditional variance forecasts generated by each of the J models for that period.

Schematically, the algorithm used to obtain the combined predictor *EN-Comb* is described as follows, omitting the low-frequency time index t for the sake of simplicity:

1. **Estimate** all the J models over a window of the *training* sample T_{in}, including observations from

the day $i = p+1$ to $i = p+T_{in}$, with starting value $p = 0$.

2. Conditionally on the estimated parameters, **generate** H one-step ahead conditional variance forecasts, for all the J models.

3. **Estimate** $\widehat{\boldsymbol{\beta}}$ via the EN approach in (9) over the *training* sample.

4. **Obtain** EN-Comb for the period $i = (p+T_{in}+1)$ to $i = (p+T_{in}+H)$, combining each of the H one-step ahead conditional variance forecasts of the models through $\widehat{\boldsymbol{\beta}}$ of the previous step via Eq. (10).

5. Moving p ahead by H, that is, $p = \{H, 2H, \cdots, (nstep-1)H\}$, **repeat** steps 1, 2, 3, and 4, $(nstep-1)$ times.

3 Empirical Analysis

3.1 The dataset

We employ daily stock log-returns and eleven variants of monthly EPUs for the 50 US states to forecast volatility using the GARCH-MIDAS model, covering the period of January, 2010 to December, 2023.[1] The state-level equity indices are taken from the Bloomberg terminal as the capitalization-weighted index of equities domiciled in a given state.[2]

In the conditional heteroscedasticity framework, the presence of outliers may alter the GARCH estimates (Carnero et al., 2007). Therefore, we use the procedure suggested by Hoaglin and Iglewicz (1987) to detect outliers in log-returns, possibly induced by thin trading. On average, 12.7 outliers per state are found. When an outlier is detected, it is replaced by the overall median. Washington has the lowest number of detected outliers (6 outliers), while Hawaii has the largest number (19 outliers).

In the first two columns of Table 1, for each US state in decreasing order by population size, we report the number of stocks included in the state-level equity index and the total market capitalization of these stocks (in $ '000). This gives a good idea of the basis of the heterogeneity at stake. The remaining columns of Table 1 include the summary statistics of the (outlier-free) log-returns (in percentage annualized terms) confirms the heterogeneity of market outcomes across states: while many states experience average daily returns close to zero, the range extends from notable losses in Wyoming (an annualized value of -2.04%) to gains in Vermont (an annualized value of 1.49%); overall volatility represented by the standard deviation of the log-returns ranges between 14.2% (New Jersey) and 50% in Wyoming, with most states between 15% and 26%. Skewness is generally negative (40 out of 50) and excess kurtosis is noticeable but not excessive (generally under 3, with rare exceptions). The descriptive picture testifies to the unique risk-return profiles across states, emphasizing the importance of considering state-specific factors in financial modeling and investment decisions. Graphically, and limiting it to the first ten US states by population for space constraints, the time series of the log-returns are plotted in Figure 1,[3] where the usual episodes of heightened volatility are discernible, with the COVID-19 crisis in 2020 the most evident one with corresponding sudden spikes and larger fluctuations of the log-returns from there on.

[1] The EPU indices are uniformly available for all states in 2006; since we need to skip 24 monthly observations for the MIDAS filter to be operational, we decided to further move forward the start of the estimation of the models away from the Lehman Brothers demise in September 2008.

[2] This approach is motivated by the core business activities of firms often taking place near their headquarters, influenced by the economic and financial dynamics of that particular state.

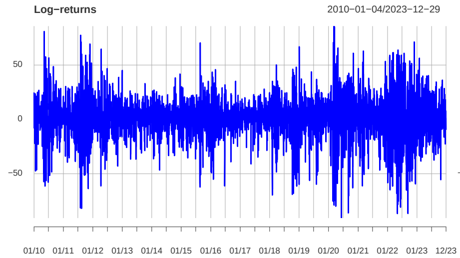
[3] The remaining forty plots are reported in the supplementary material.

Table 1: Summary statistics by state in decreasing order by population.

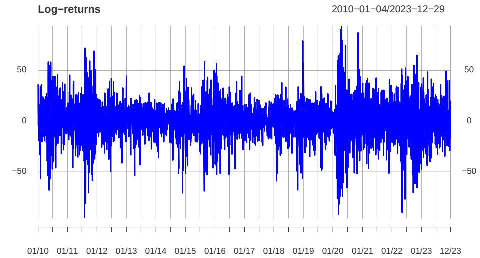
State	# Equities	Capitalization	Minimum	Maximum	Mean	Std. Dev.	Skewness	Kurtosis
California	972	14210.613	−90.654	85.671	0.859	19.763	−0.402	2.400
Texas	492	4173.075	−95.643	93.786	0.395	19.074	−0.204	2.387
Florida	292	853.422	−81.646	82.718	0.686	16.951	−0.246	2.444
New York	638	4462.509	−84.040	78.294	0.526	16.230	−0.310	3.002
Pennsylvania	186	1022.095	−78.388	75.247	0.586	16.809	−0.289	2.284
Illinois	175	2013.918	−82.558	73.244	0.661	15.110	−0.364	3.026
Ohio	128	1466.353	−77.452	73.455	0.697	14.613	−0.171	3.089
Georgia	127	1340.082	−82.431	71.390	0.674	15.038	−0.386	2.742
North Carolina	87	999.935	−95.411	94.048	0.759	19.481	−0.170	2.544
Michigan	71	470.490	−89.839	94.764	0.691	19.317	−0.170	2.304
New Jersey	178	1345.889	−71.914	66.841	0.493	14.158	−0.356	2.655
Virginia	127	838.467	−78.848	70.921	0.701	15.387	−0.254	2.786
Washington	100	5341.921	−100.565	111.501	1.171	21.827	−0.151	2.869
Arizona	83	475.685	−101.438	105.969	0.643	21.671	−0.046	1.902
Tennessee	57	374.653	−82.259	84.284	0.891	17.730	−0.117	2.216
Massachusetts	308	1784.901	−89.144	84.799	0.651	18.090	−0.331	2.318
Indiana	60	969.007	−81.469	88.445	0.878	16.702	−0.136	2.300
Missouri	42	305.228	−85.162	80.788	0.645	16.214	−0.139	2.744
Maryland	91	375.882	−88.276	86.590	0.571	17.238	−0.253	2.714
Wisconsin	60	325.679	−97.391	92.929	0.726	18.968	−0.158	2.615
Colorado	157	314.641	−93.673	84.483	0.415	18.030	−0.271	2.221
Minnesota	81	1059.701	−82.781	66.364	0.673	15.453	−0.300	2.296
South Carolina	22	22.331	−90.008	110.406	0.578	20.775	0.232	2.659
Alabama	19	67.972	−119.981	116.727	0.712	22.872	−0.073	2.652
Louisiana	22	59.502	−94.445	89.067	0.357	17.987	−0.122	2.629
Kentucky	24	158.106	−90.327	83.812	0.948	17.030	−0.251	2.976
Oregon	24	161.893	−112.857	107.901	0.648	22.060	−0.212	2.995
Oklahoma	33	161.761	−141.016	131.172	0.523	26.366	−0.109	3.166
Connecticut	82	850.044	−92.416	88.063	0.792	18.253	−0.319	2.689
Utah	50	71.018	−90.209	98.344	0.700	20.340	−0.203	1.834
Iowa	18	51.526	−96.643	106.641	0.892	21.350	−0.156	2.415
Nevada	103	132.874	−149.123	146.607	0.758	28.661	0.020	2.748
Arkansas	17	500.597	−76.972	80.845	0.402	15.544	−0.241	2.507
Kansas	20	18.835	−132.017	129.751	0.324	26.552	−0.027	2.867
Mississippi	10	24.111	−104.761	106.683	0.539	21.204	0.064	2.431
New Mexico	5	6.208	−174.778	206.207	0.489	32.385	0.536	6.635
Nebraska	17	952.964	−89.681	85.207	0.917	17.374	0.030	2.763
Idaho	15	133.883	−149.888	134.659	1.066	33.283	−0.107	1.569
West Virginia	8	9.266	−118.863	132.613	0.470	25.499	0.154	2.916
Hawaii	13	14.115	−94.652	108.519	0.488	20.495	0.100	2.685
New Hampshire	11	22.384	−100.733	100.645	0.729	20.473	−0.097	2.345
Maine	7	56.577	−114.223	119.137	1.244	24.352	−0.146	2.382
Montana	7	8.161	−145.040	154.186	0.583	30.246	0.040	2.499
Rhode Island	14	146.885	−102.689	84.820	0.630	19.812	−0.244	2.036
Delaware	41	148.256	−111.694	111.382	0.452	23.396	−0.110	1.872
South Dakota	5	8.274	−97.314	109.634	0.519	19.964	−0.052	2.788
North Dakota	8	9.775	−114.056	101.378	0.253	21.068	−0.275	2.641
Alaska	1	0.293	−107.250	104.048	0.641	21.200	−0.030	2.659
Vermont	5	5.036	−188.559	231.310	1.498	36.217	0.417	5.859
Wyoming	8	0.706	−255.677	278.609	−2.042	49.905	0.288	3.900

Notes: The table reports the number of stocks included in the state-level equity index and the total market capitalization of these stocks (in \$ '000) in the first two columns. The remaining columns present summary statistics of the (outlier-free) daily-log returns in annualized percentage terms. Kurtosis is the excess kurtosis. Sample period: January 2010 to December 2023 (3531 observations).

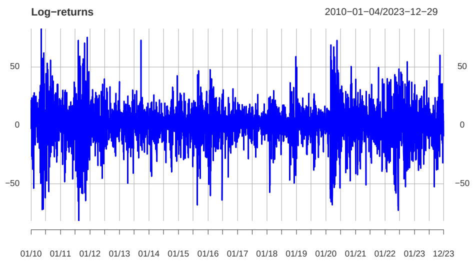
Figure 1: Daily log-returns for the ten most populated states, expressed in annualized percentage terms (different scales).



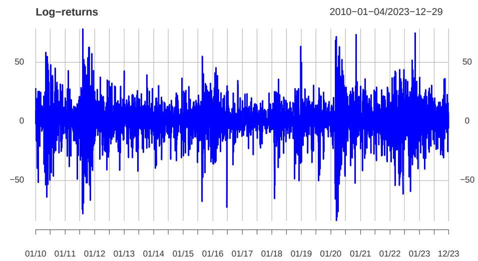
(a) California



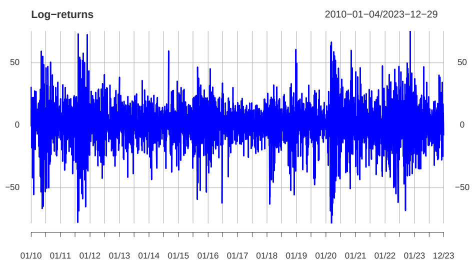
(b) Texas



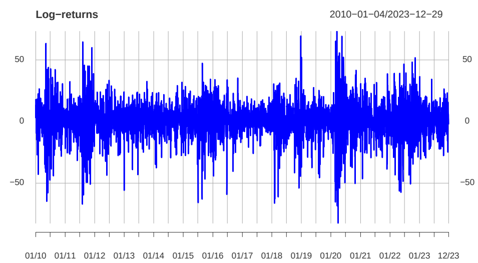
(c) Florida



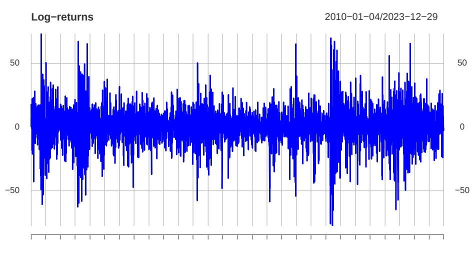
(d) New York



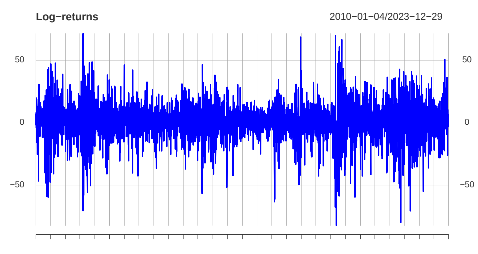
(e) Pennsylvania



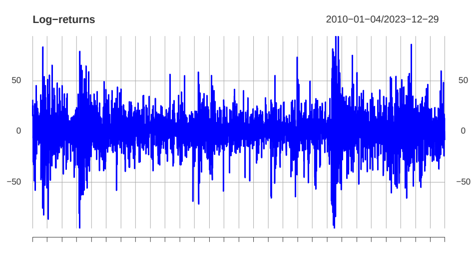
(f) Illinois



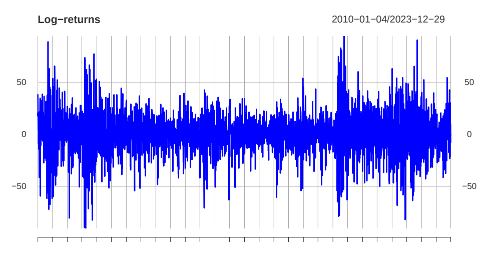
(g) Ohio



(h) Georgia



(i) North Carolina



(j) Michigan

The monthly EPU indices used in this work, as developed by Baker et al. (2022), are derived from the site https://policyuncertainty.com/state_epu.html, which provides the *National*, *Local*, and *Composite* EPU indices for each US state. To construct the state-level measures of EPU, Baker et al. (2022) use around 3,500 daily and weekly newspapers for every state in the US (as well as Washington DC), but exclude national papers published in a given state (such as the New York Times or the Wall Street Journal). The three state-level EPU indices are constructed by recording the fraction of articles that contain terms from sets regarding the economy, uncertainty, and policy: the nation-level EPU index measures the level of uncertainty within a state that stems from policy-related sources with a national content.[4] By contrast, the state-level index aims at capturing uncertainty within a state, stemming from state and local policy issues.[5] Finally, the third composite index is derived from articles that contain terms related to the economy and uncertainty and a term from a set containing both state-specific and national policy terms. These three EPU indices are summarized in the first three rows of Table 2 and illustrated in Figure 2 (for the first ten US states by population), while all the remaining plots are in the supplementary material.

With the idea to enlarge the set of low-frequency variables to capture extra elements in the measurement of uncertainty in policy issues, we computed seven other novel indices derived as averages of the existing EPU indices, detailed in Table 2; they take into consideration a national aggregation of the EPU indices and also variants at the state level that averages across neighboring states as well. Moreover, and this brings the number of low-frequency variables to eleven, Baker et al. (2016) also provide a global US EPU index, available at: https://policyuncertainty.com/us_monthly.html, based on news coverage with a weighting scheme on the broad news-based policy uncertainty index, the tax expiration indices, the CPI forecast disagreement measure, and the federal/state/local purchases disagreement metric.

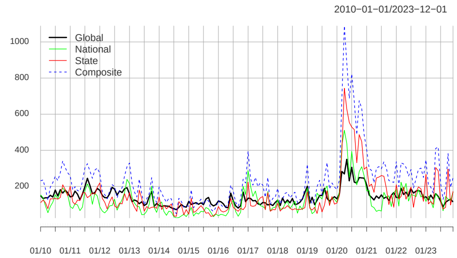
Table 2: Low-frequency variables.

Label	Full Name	Level	Informative Basis
Nat.	National	State	National policy-related sources
Loc.	Local	State	State and local policy issues
Comp.	Composite	State and National	National and state-specific EPU sources
Avg.	EPU Averaged	State	Average of state-specific Nat., Loc. and Comp. indexes
Nat. Avg.	National Averaged	National	Average of all the Nat. indexes
Loc. Avg.	Local Averaged	National	EPU Average of all the Loc. indexes
Comp. Avg.	Composite Averaged	National	Average of all the Comp. indexes
Neigh. Nat.	Neighbourhood National	State	Average of all the Nat. indexes of the bordering states
Neigh. Loc.	Neighbourhood Local	State	Average of all the Loc. indexes of the bordering states
Neigh. Comp.	Neighbourhood Composite	State	Average of all the Comp. indexes of the bordering states
Glob.	Global	National	Global (at US level) EPU

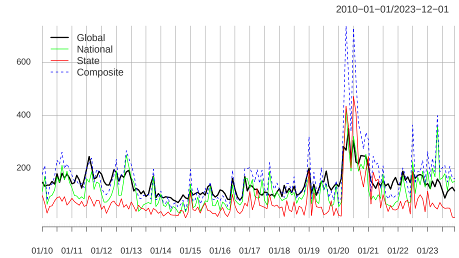
[4] It includes terms related to national elections, elected officials, federal agencies, departments, and regulators.

[5] Each state-specific policy term set includes terms that describe the names of their executive positions and legislative bodies at both state and local levels as well as terms that note policy initiatives put to a direct vote by citizens. Also included are the names of the state bodies that deal with regulations spanning the environment, labor and unemployment, gambling, transportation, banking, energy and utilities, and other financial services. As a result, this set of terms is unique to each state, since the names and titles of officials and regulators and departments vary across states.

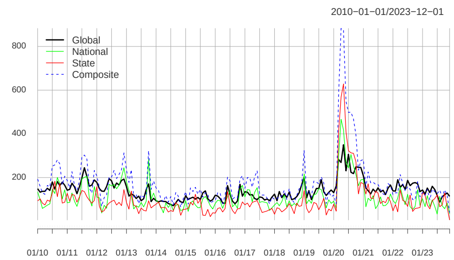
Figure 2: Global and Three State-specific EPUs.



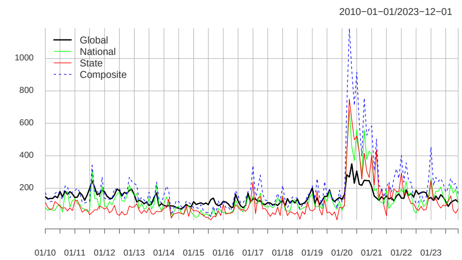
(a) California



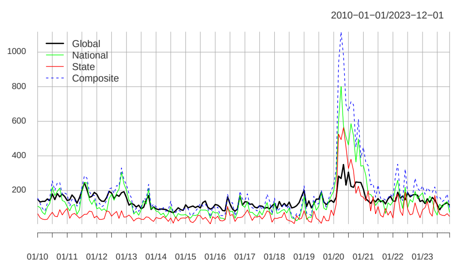
(b) Texas



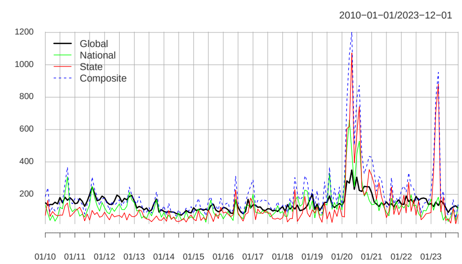
(c) Florida



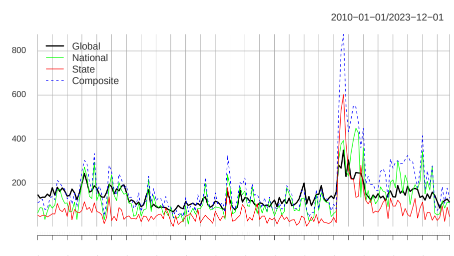
(d) New York



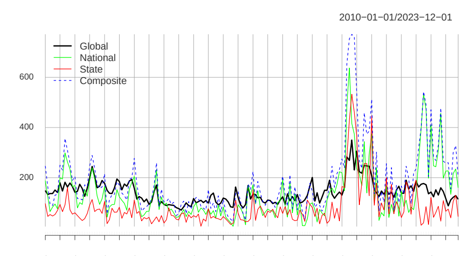
(e) Pennsylvania



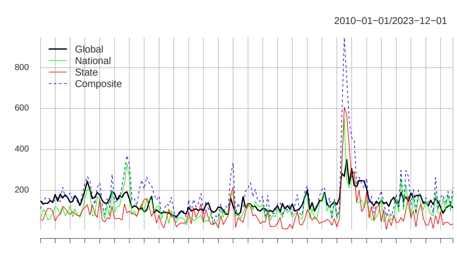
(f) Illinois



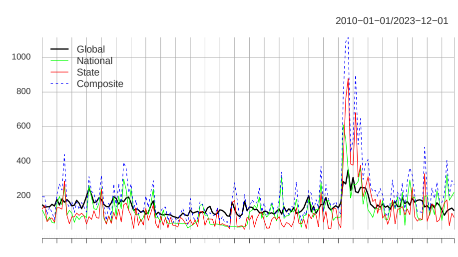
(g) Ohio



(h) Georgia



(i) North Carolina



(j) Michigan

3.2 Estimation results

Overall, we estimate eleven GARCH-MIDAS models, plus the GJR specification. We adopt a rolling window of eleven years, which, according to the notation of the previous section, means that T_{in}, the length of the rolling window, is equal to 2771 daily observations. The out-of-sample period covers the time span from 2021 to 2023, corresponding to 750 daily observations. Moreover, we initially consider the length of the forecasting horizon, H, to be 30, meaning that after each estimation of the training sample, the conditional variance forecasts are generated for the following thirty days. Finally, the combined predictor, EN-Comb, is adaptively obtained by combining the conditional variance forecasts of the models through Eq. (10). The frequency of model re-estimation (and then of the EN-comb) is, again, fixed to 30 days.

Table 3 reports the estimated coefficients for the first training period of the *theta*s of the eleven GARCH-MIDAS models (for space constraints, all the other variables are available upon request). Dark, medium and light shades of green denote the significance at levels $1\%, 5\%, 10\%$, respectively. In line with the theories outlined above, the effect of EPUs on stock returns volatility is mixed, with the significant effects being both positive and negative. But evidently, irrespective of the direction of the impact, clearly there is strong prrof that EPUs do predict stock market volatility of the 50 US states.

In order to implement the EN-Comb predictor, we need to replace $\sigma_{i,t}^2$ with a (unbiased) volatility proxy. In absence of high-frequency data, we resort to use the squared daily log-returns as (unbiased, even if noisy) volatility proxy. Once run the EN regression for each training period and once obtained the EN-Comb predictor, we jointly evaluate the conditional variance of EN-Comb, together that all the other conditional variances from the other twelve models, via the MCS. As in Cipollini et al. (2021) and references therein, the test statistic used in the MCS is the semi-quadratic, denoted by T_{SQ}:

$$T_{SQ} = \sum_{\mathrm{i} \neq \mathrm{j} \in J} \frac{\bar{\ell}_{\mathrm{i,j}}^2}{\widehat{var}\left(\bar{\ell}_{\mathrm{i,j}}\right)},$$

where i and j denote two models within the model universe J, $\bar{\ell}$ the loss differential between i and j, and $\widehat{var}\left(\bar{\ell}_{\mathrm{i,j}}\right)$ the variance of $\bar{\ell}_{\mathrm{i,j}}$ calculated using a bootstrap procedure on 5,000 replications.

Table 4 reports the Mean Squared Errors (MSE) of the twelve models involved in our analysis and the combined predictor EN-Comb, for the out-of-sample period. Green cells indicate inclusion in the SSM at 25% significance level. Remarkably, the combined predictor, based on the low-frequency variables concerning the local and global EPUs, almost always (more precisely, 45 times out of 50) enters the SSM of the MCS. Moreover, in some states, such as New York, Georgia, and Connecticut, EN-Comb is the only model included in the SSM. Furthermore, when the SSM is large, the MSE of EN-Comb is usually relatively smaller than that of the others (see, for instance, the cases of New Hampshire, Rhode Island, and North Dakota).

The estimated $\boldsymbol{\beta}$'s of the *training* samples obtained through Eq. (9) and employed in Eq. (10) may be positive or negative. Therefore, it is worth verifying the robustness of the results when positivity constraints are imposed on the $\boldsymbol{\beta}$'s. Table 5 reports the MSE of the whole model universe, with the EN-Comb predictor obtained using such a positivity constraint in Eq. (9). Interestingly, the results found in Table 4 are largely confirmed: (i) the EN-Comb predictor enters the SSM always, except three times; (ii) EN-Comb is the only model belonging to the SSM for three states, namely Georgia, Oregon, and Connecticut.

A detailed robustness analysis of the results is not reported here for space limits but it is reported within the supplementary material. In particular, we repeat the analysis using the QLIKE loss function and updating more frequently the EN-Comb combined predictor, with the positivity constraints and without. Overall, the results presented are by and large confirmed.

Table 3: In-sample estimates of the θ parameter of the eleven GARCH-MIDAS models.

State/EPU	Nat.	Loc.	Comp.	Avg.	*Nat. Avg.*	*Loc. Avg.*	*Comp. Avg.*	*Neigh. Nat.*	*Neigh. Loc.*	*Neigh. Comp.*	Glob.
California	−0.185	3.034	−0.304	2.853	2.641	1.576	3.331	1.194	2.815	4.987	−0.723
Texas	−0.14	0.008	−0.459	−0.123	−0.486	0.579	0.442	−0.215	0.668	−0.277	−0.483
Florida	−0.228	0.733	−0.015	8.496	−3.214	−1.068	−0.498	−1.719	−1.202	−2.396	−7.921
New York	−0.092	0.265	0.085	0.779	−2.085	−0.85	−2.088	−2.106	1.793	0.407	−1.661
Pennsylvania	−1.809	−0.148	−1.621	−0.69	−0.445	0.266	−0.62	−0.353	0.243	−0.287	−3.573
Illinois	−0.464	0.261	−0.557	−0.362	−0.949	0.731	−0.16	−1.359	0.57	−0.141	−1.091
Ohio	−0.257	0.579	−0.267	0.281	−0.981	0.733	0.732	0.301	0.855	−0.725	−0.732
Georgia	−0.733	−0.346	−1.714	−1.103	0.02	1.058	1.554	2.612	1.67	1.99	1.173
North Carolina	1.407	−0.142	−0.428	−0.194	−0.177	−0.244	0.143	0.83	1.001	0.8	4.716
Michigan	0.892	−0.505	−0.901	−0.981	−2.243	−1.188	0.018	−0.256	−1.698	−1.934	−3.419
New Jersey	−4.281	−0.187	−1.277	−0.491	−0.592	0.414	−0.152	−1.368	0.343	−0.34	−7.204
Virginia	−0.147	0.036	0.438	−0.082	−1.518	0.66	−0.791	−1.995	0.644	−1.211	−4.368
Washington	2.096	1.211	2.817	2.223	0.941	1.112	2.229	1.546	1.235	1.734	−2.589
Arizona	0.623	0.74	1.814	1.675	−1.717	0.428	−0.178	−1.386	0.131	0.524	−2.704
Tennessee	0.247	−0.46	−0.888	0.22	−0.198	0.801	0.648	1.386	1.485	1.634	−0.879
Massachusetts	−0.728	0.889	−0.111	−0.257	−0.648	0.658	−0.387	−0.307	1.181	−0.232	−1.599
Indiana	−0.775	−0.143	−0.992	−0.556	−1.492	0.55	−0.582	−2.047	0.898	−0.274	−1.164
Missouri	−0.103	0.058	−0.087	0.16	−2.512	0.203	−2.374	−0.151	0.835	−0.202	−0.241
Maryland	−1.058	0.106	−2.222	−1.453	−2.029	−0.548	−0.248	2.389	4.671	1.72	−4.627
Wisconsin	−3.128	−1.485	−1.732	−2.458	−0.933	−1.657	−0.301	−2.626	0.318	−2.228	−5.874
Colorado	0.04	0.115	0.152	0.086	−0.947	0.289	−2.134	0.14	0.937	0.522	−5.373
Minnesota	−1.089	−0.074	0.097	−1.389	−0.228	0.625	0.449	0.103	1.076	1.636	−0.971
South Carolina	−1.082	0.115	0.457	0.499	0.121	0.995	0.495	0.671	0.08	1.014	0.659
Alabama	−0.328	0.046	−0.2	0.089	−3.036	0.308	−1.944	−3.483	−0.015	−0.325	−6.898
Louisiana	0.365	−1.125	−0.515	0.395	1.018	0.741	1.23	−0.365	1.548	−0.284	4.333
Kentucky	−0.278	0.19	0.241	0.182	0.363	0.705	0.835	−0.224	0.739	−0.108	−0.415
Oregon	1.064	1.083	−0.091	1.456	2.269	1.147	2.913	2.066	1.361	2.1	5.882
Oklahoma	0.168	0.203	0.009	0.127	2.914	0.887	1.076	−0.173	2.772	0.386	1.096
Connecticut	−1.267	0.354	−0.756	−0.772	−2.83	−1.644	−3.609	−0.217	−1.643	−2.208	−5.364
Utah	−0.542	0.22	−0.604	0.365	−1.381	0.49	0.321	0.434	−0.009	1.286	−2.541
Iowa	−1.539	0.071	−1.609	0.11	−2.49	−1.227	−2.909	−0.635	−3.802	−3.384	−7.805
Nevada	0.079	1.151	−0.104	0.553	0.37	0.215	0.828	1.582	−2.486	1.481	2.099
Arkansas	−0.507	0.101	−0.25	−0.142	0.419	0.287	−0.055	−0.149	0.357	0.322	−0.249
Kansas	−0.024	0.069	0.659	0.098	1.271	0.43	2.047	−1.127	3.763	0.677	9.06
Mississippi	−0.629	0.732	−0.237	−0.209	−0.036	0.881	1.108	3.762	0.097	1.7	−6.961
New Mexico	0.251	−0.241	2.025	0.754	2.61	2.457	6.672	0.291	5.368	3.524	14.168
Nebraska	0.000	0.657	1.319	1.077	−0.039	0.848	2.171	2.217	1.932	2.588	1.806
Idaho	0.169	0.162	0.524	0.159	−0.296	0.689	1.477	1.723	0.252	1.352	3.656
West Virginia	0.211	0.379	0.733	0.184	−0.501	1.896	0.233	−0.571	2.202	−0.236	−4.132
Hawaii	0.814	0.795	−1.043	0.849	−1.641	0.278	−0.741				−8.114
New Hampshire	−0.058	−0.02	−0.418	−0.049	−1.231	0.295	−0.653	−0.197	0.346	−1.065	−8.033
Maine	−0.828	0.511	−0.794	−0.475	−0.286	0.579	−0.247	−0.035	2.658	0.847	−0.795
Montana	−0.241	−0.618	−2.291	−1.586	−2.523	0.098	−2.919	−1.456	−2.468	−3.039	−3.303
Rhode Island	0.183	0.385	0.896	0.47	2.076	1.565	3.138	−1.102	2.084	−0.14	1.77
Delaware	−0.064	−0.004	0.095	0.067	−0.193	−0.058	−0.17	−0.209	−0.177	−0.257	−0.755
South Dakota	−0.877	0.13	−1.063	−1.095	−2.371	0.251	−1.624	−0.207	0.102	−1.108	−2.528
North Dakota	0.474	0.445	2.098	1.091	−0.884	0.887	−0.08	−1.568	1.76	−1.392	0.583
Alaska	−1.156	−0.435	−1.782	−0.28	3.073	1.622	3.512				12.554
Vermont	−0.496	0.253	0.028	0.094	0.07	0.33	0.312	0.353	2.666	2.104	−0.719
Wyoming	0.036	0.685	−0.004	0.258	−1.287	−1.338	−1.868	−2.029	−2.839	−2.456	10.472

Notes: For an explanation of the low-frequency variables, see Table 2. Dark, medium, and light shades of green signal the coefficient significance at 1%, 5%, 10% levels, respectively. Quasi-Maximum Likelihood (Bollerslev and Wooldridge, 1992) standard errors have been used. Sample period: January 2010 to December 2020 (2771 observations).

Table 4: Out-of-sample volatility evaluation. MSE loss function. EN-Comb updated every 30 days.

State/EPU	GARCH-MIDAS											GJR	EN-Comb
	Nat.	Loc.	Comp.	Avg.	Nat. Avg.	Loc. Avg.	Comp. Avg.	Neigh. Nat.	Neigh. Loc.	Neigh. Comp.	Glob.		
California	0.943	0.98	0.952	0.945	0.963	0.969	0.975	0.947	0.985	0.985	0.948	0.9	0.892
Texas	0.487	0.473	0.474	0.476	0.474	0.475	0.475	0.474	0.475	0.474	0.475	0.471	0.472
Florida	0.306	0.308	0.306	0.306	0.306	0.307	0.306	0.306	0.305	0.305	0.306	0.3	0.301
New York	0.236	0.235	0.236	0.236	0.236	0.235	0.236	0.248	0.235	0.235	0.238	0.232	0.227
Pennsylvania	0.295	0.289	0.29	0.29	0.29	0.291	0.29	0.289	0.29	0.29	0.291	0.284	0.285
Illinois	0.155	0.155	0.157	0.153	0.155	0.154	0.155	0.155	0.155	0.154	0.157	0.152	0.152
Ohio	0.175	0.188	0.177	0.177	0.177	0.176	0.177	0.177	0.176	0.177	0.177	0.174	0.174
Georgia	0.257	0.259	0.258	0.257	0.259	0.262	0.261	0.261	0.262	0.262	0.258	0.255	0.248
North Carolina	0.458	0.458	0.459	0.458	0.457	0.458	0.474	0.458	0.458	0.459	0.458	0.453	0.467
Michigan	0.58	0.579	0.579	0.579	0.58	0.582	0.58	0.578	0.577	0.58	0.579	0.573	0.576
New Jersey	0.086	0.086	0.085	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.087	0.085	0.085
Virginia	0.205	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.203	0.203	0.201	0.2
Washington	1.199	1.217	1.215	1.222	1.208	1.206	1.195	1.182	1.195	1.201	1.189	1.164	1.176
Arizona	0.917	0.917	0.922	0.916	0.918	0.911	0.919	0.917	0.921	0.917	0.915	0.896	0.902
Tennessee	0.286	0.286	0.286	0.286	0.286	0.286	0.286	0.292	0.29	0.287	0.287	0.286	0.285
Massachusetts	0.387	0.43	0.392	0.386	0.387	0.388	0.389	0.389	0.389	0.389	0.391	0.375	0.373
Indiana	0.334	0.333	0.335	0.334	0.335	0.338	0.335	0.335	0.337	0.336	0.335	0.332	0.335
Missouri	0.212	0.212	0.212	0.213	0.212	0.211	0.212	0.212	0.213	0.212	0.213	0.214	0.212
Maryland	0.254	0.25	0.247	0.248	0.248	0.249	0.248	0.248	0.248	0.249	0.249	0.243	0.244
Wisconsin	0.4	0.395	0.399	0.399	0.398	0.398	0.397	0.398	0.411	0.397	0.403	0.392	0.399
Colorado	0.5	0.499	0.5	0.498	0.5	0.502	0.501	0.502	0.505	0.503	0.501	0.492	0.489
Minnesota	0.198	0.198	0.198	0.197	0.198	0.199	0.199	0.198	0.197	0.198	0.199	0.197	0.195
South Carolina	0.998	0.996	0.999	0.998	0.998	0.996	0.997	0.996	0.994	0.996	0.998	0.994	1.014
Alabama	0.678	0.678	0.679	0.68	0.681	0.677	0.679	0.681	0.722	0.68	0.686	0.68	0.678
Louisiana	0.553	0.549	0.553	0.553	0.553	0.554	0.555	0.55	0.555	0.551	0.553	0.549	0.552
Kentucky	0.294	0.295	0.294	0.294	0.293	0.294	0.294	0.293	0.294	0.293	0.294	0.29	0.289
Oregon	1.627	1.65	1.763	1.628	1.647	1.649	1.655	1.653	1.657	1.656	1.64	1.611	1.604
Oklahoma	1.355	1.349	1.352	1.358	1.355	1.356	1.354	1.356	1.357	1.361	1.354	1.359	1.355
Connecticut	0.394	0.39	0.392	0.392	0.395	0.394	0.393	0.391	0.391	0.393	0.398	0.388	0.384
Utah	0.693	0.693	0.694	0.693	0.694	0.692	0.696	0.695	0.699	0.7	0.688	0.684	0.685
Iowa	0.545	0.543	0.543	0.543	0.545	0.544	0.546	0.545	0.546	0.546	0.545	0.545	0.545
Nevada	2.092	2.102	2.091	2.096	2.085	2.089	2.087	2.106	2.095	2.105	2.085	2.07	2.088
Arkansas	0.164	0.164	0.164	0.164	0.164	0.164	0.164	0.164	0.165	0.164	0.164	0.164	0.164
Kansas	4.046	4.047	4.038	4.056	4.044	4.045	4.041	4.062	4.023	4.061	4.074	4.032	4.031
Mississippi	0.766	0.769	0.77	0.767	0.768	0.769	0.769	0.768	0.769	0.765	0.771	0.767	0.776
New Mexico	16.607	16.209	19.155	20.96	22.668	16.358	16.735	16.229	16.243	17.012	16.233	16.388	17.327
Nebraska	0.242	0.242	0.244	0.244	0.244	0.243	0.246	0.243	0.244	0.244	0.243	0.241	0.242
Idaho	2.335	2.337	2.332	2.354	2.336	2.334	2.338	2.32	2.308	2.316	2.305	2.322	2.334
West Virginia	1.713	1.711	1.718	1.695	1.694	1.707	1.708	1.711	1.707	1.707	1.716	1.689	1.761
Hawaii	1.156	1.152	1.161	1.154	1.156	1.157	1.157				1.157	1.147	1.162
New Hampshire	0.759	0.756	0.76	0.76	0.759	0.758	0.758	0.76	0.759	0.759	0.763	0.759	0.757
Maine	2.105	2.14	2.118	2.115	2.126	2.132	2.104	2.114	2.101	2.151	2.121	2.063	2.089
Montana	3.295	3.259	3.28	3.268	3.278	3.285	3.213	3.264	3.224	3.257	3.285	3.217	3.311
Rhode Island	0.631	0.633	0.63	0.632	0.633	0.635	0.634	0.63	0.632	0.63	0.63	0.63	0.624
Delaware	1.295	1.294	1.29	1.295	1.294	1.294	1.294	1.295	1.297	1.285	1.296	1.273	1.275
South Dakota	0.579	0.578	0.577	0.579	0.579	0.578	0.578	0.575	0.577	0.582	0.577	0.576	0.578
North Dakota	0.636	0.634	0.635	0.634	0.634	0.63	0.632	0.632	0.632	0.632	0.631	0.632	0.629
Alaska	1.208	1.215	1.213	1.214	1.221	1.213	1.213				1.215	1.208	1.223
Vermont	4.132	4.083	4.08	4.153	4.079	4.079	4.08	4.154	4.003	4.026	4.131	4.218	4.823
Wyoming	37.293	37.46	37.206	37.354	37.437	36.959	37.09	37.098	37.352	37.138	39.634	37.441	37.959

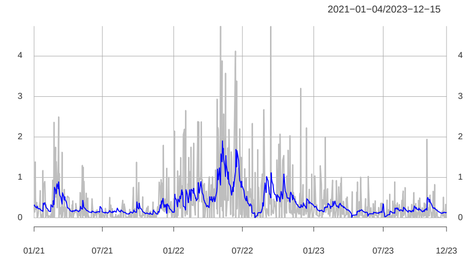
Notes: The table reports the MSE (multiplied by 100) of the GARCH-MIDAS models (columns from two to the antepenultimate), the GJR model, and the proposed predictor EN-Comb. The low-frequency variables of the GARCH-MIDAS models in columns from two to penultimate are illustrated in Table 2. Sample period: January 2021 to December 2023 (750 observations). Green cells indicate inclusion in the SSM at 25% significance level.

Table 5: Out-of-sample volatility evaluation. MSE loss function. EN-Comb updated every 30 days, with $\widehat{\beta} \geq 0$.

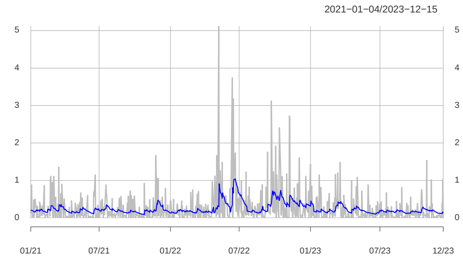
State/EPU	GARCH-MIDAS											GJR	EN-Comb
	Nat.	Loc.	Comp.	Avg.	Nat. Avg.	Loc. Avg.	Comp. Avg.	Neigh. Nat.	Neigh. Loc.	Neigh. Comp.	Glob.		
California	0.943	0.98	0.952	0.945	0.963	0.969	0.975	0.947	0.985	0.985	0.948	0.9	0.896
Texas	0.487	0.473	0.474	0.476	0.474	0.475	0.475	0.474	0.475	0.474	0.475	0.471	0.472
Florida	0.306	0.308	0.306	0.306	0.306	0.307	0.306	0.306	0.305	0.305	0.306	0.3	0.298
New York	0.236	0.235	0.236	0.236	0.236	0.235	0.236	0.248	0.235	0.235	0.238	0.232	0.228
Pennsylvania	0.295	0.289	0.29	0.29	0.29	0.291	0.29	0.289	0.29	0.29	0.291	0.284	0.285
Illinois	0.155	0.155	0.157	0.153	0.155	0.154	0.155	0.155	0.155	0.154	0.157	0.152	0.15
Ohio	0.175	0.188	0.177	0.177	0.177	0.176	0.177	0.177	0.176	0.177	0.177	0.174	0.175
Georgia	0.257	0.259	0.258	0.257	0.259	0.262	0.261	0.261	0.262	0.262	0.258	0.255	0.248
North Carolina	0.458	0.458	0.459	0.458	0.457	0.458	0.474	0.458	0.458	0.459	0.458	0.453	0.459
Michigan	0.58	0.579	0.579	0.579	0.58	0.582	0.58	0.578	0.577	0.58	0.579	0.573	0.576
New Jersey	0.086	0.086	0.085	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.087	0.085	0.085
Virginia	0.205	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.203	0.203	0.201	0.2
Washington	1.199	1.217	1.215	1.222	1.208	1.206	1.195	1.182	1.195	1.201	1.189	1.164	1.178
Arizona	0.917	0.917	0.922	0.916	0.918	0.911	0.919	0.917	0.921	0.917	0.915	0.896	0.907
Tennessee	0.286	0.286	0.286	0.286	0.286	0.286	0.286	0.292	0.29	0.287	0.287	0.286	0.285
Massachusetts	0.387	0.43	0.392	0.386	0.387	0.388	0.389	0.389	0.389	0.389	0.391	0.375	0.378
Indiana	0.334	0.333	0.335	0.334	0.335	0.338	0.335	0.335	0.337	0.336	0.335	0.332	0.335
Missouri	0.212	0.212	0.212	0.213	0.212	0.211	0.212	0.212	0.213	0.212	0.213	0.214	0.211
Maryland	0.254	0.25	0.247	0.248	0.248	0.249	0.248	0.248	0.248	0.249	0.249	0.243	0.244
Wisconsin	0.4	0.395	0.399	0.399	0.398	0.398	0.397	0.398	0.411	0.397	0.403	0.392	0.39
Colorado	0.5	0.499	0.5	0.498	0.5	0.502	0.501	0.502	0.505	0.503	0.501	0.492	0.492
Minnesota	0.198	0.198	0.198	0.197	0.198	0.199	0.199	0.198	0.197	0.198	0.199	0.197	0.195
South Carolina	0.998	0.996	0.999	0.998	0.998	0.996	0.997	0.996	0.994	0.996	0.998	0.994	1.003
Alabama	0.678	0.678	0.679	0.68	0.681	0.677	0.679	0.681	0.722	0.68	0.686	0.68	0.675
Louisiana	0.553	0.549	0.553	0.553	0.553	0.554	0.555	0.55	0.555	0.551	0.553	0.549	0.552
Kentucky	0.294	0.295	0.294	0.294	0.293	0.294	0.294	0.293	0.294	0.293	0.294	0.29	0.289
Oregon	1.627	1.65	1.763	1.628	1.647	1.649	1.655	1.653	1.657	1.656	1.64	1.611	1.585
Oklahoma	1.355	1.349	1.352	1.358	1.355	1.356	1.354	1.356	1.357	1.361	1.354	1.359	1.352
Connecticut	0.394	0.39	0.392	0.392	0.395	0.394	0.393	0.391	0.391	0.393	0.398	0.388	0.384
Utah	0.693	0.693	0.694	0.693	0.694	0.692	0.696	0.695	0.699	0.7	0.688	0.684	0.687
Iowa	0.545	0.543	0.543	0.543	0.545	0.544	0.546	0.545	0.546	0.546	0.545	0.545	0.545
Nevada	2.092	2.102	2.091	2.096	2.085	2.089	2.087	2.106	2.095	2.105	2.085	2.07	2.088
Arkansas	0.164	0.164	0.164	0.164	0.164	0.164	0.164	0.164	0.165	0.164	0.164	0.164	0.164
Kansas	4.046	4.047	4.038	4.056	4.044	4.045	4.041	4.062	4.023	4.061	4.074	4.032	4.028
Mississippi	0.766	0.769	0.77	0.767	0.768	0.769	0.769	0.768	0.769	0.765	0.771	0.767	0.776
New Mexico	16.607	16.209	19.155	20.96	22.668	16.358	16.735	16.229	16.243	17.012	16.233	16.388	16.787
Nebraska	0.242	0.242	0.244	0.244	0.244	0.243	0.246	0.243	0.244	0.244	0.243	0.241	0.242
Idaho	2.335	2.337	2.332	2.354	2.336	2.334	2.338	2.32	2.308	2.316	2.305	2.322	2.311
West Virginia	1.713	1.711	1.718	1.695	1.694	1.707	1.708	1.711	1.707	1.707	1.716	1.689	1.755
Hawaii	1.156	1.152	1.161	1.154	1.156	1.157	1.157				1.157	1.147	1.16
New Hampshire	0.759	0.756	0.76	0.76	0.759	0.758	0.758	0.76	0.759	0.759	0.763	0.759	0.758
Maine	2.105	2.14	2.118	2.115	2.126	2.132	2.104	2.114	2.101	2.151	2.121	2.063	2.097
Montana	3.295	3.259	3.28	3.268	3.278	3.285	3.213	3.264	3.224	3.257	3.285	3.217	3.305
Rhode Island	0.631	0.633	0.63	0.632	0.633	0.635	0.634	0.63	0.632	0.63	0.63	0.63	0.623
Delaware	1.295	1.294	1.29	1.295	1.294	1.294	1.294	1.295	1.297	1.285	1.296	1.273	1.274
South Dakota	0.579	0.578	0.577	0.579	0.579	0.578	0.578	0.575	0.577	0.582	0.577	0.576	0.578
North Dakota	0.636	0.634	0.635	0.634	0.634	0.63	0.632	0.632	0.632	0.632	0.631	0.632	0.63
Alaska	1.208	1.215	1.213	1.214	1.221	1.213	1.213				1.215	1.208	1.211
Vermont	4.132	4.083	4.08	4.153	4.079	4.079	4.08	4.154	4.003	4.026	4.131	4.218	4.142
Wyoming	37.293	37.46	37.206	37.354	37.437	36.959	37.09	37.098	37.352	37.138	39.634	37.441	37.12

Notes: The table reports the MSE (multiplied by 100) of the GARCH-MIDAS models (columns from two to the antepenultimate), the GJR model, and the proposed predictor EN-Comb. The combined predictor EN-Comb has been obtained with the constraint that the estimated *beta*s of the EN regression (Eq. (9)) are positive. The low-frequency variables of the GARCH-MIDAS models in columns from two to penultimate are illustrated in Table 2. Sample period: January 2021 to December 2023 (750 observations). Green cells indicate inclusion in the SSM at 25% significance level.

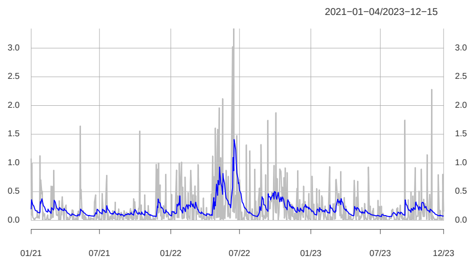
Figure 3: Out-of-sample volatility evaluation. Squared daily log-returns and EN-Comb, annualized percentage scale.



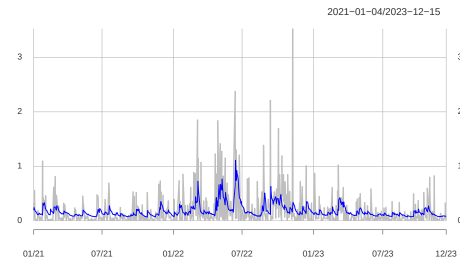
(a) California



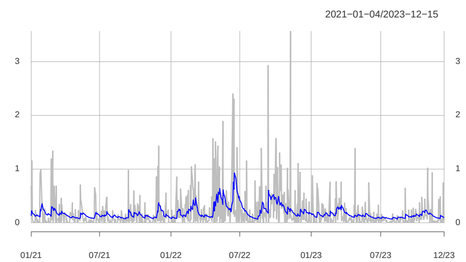
(b) Texas



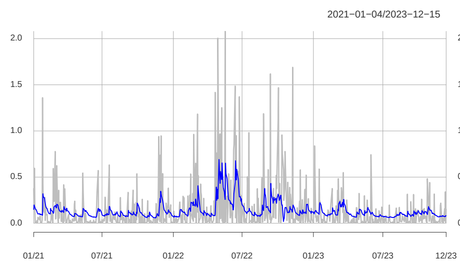
(c) Florida



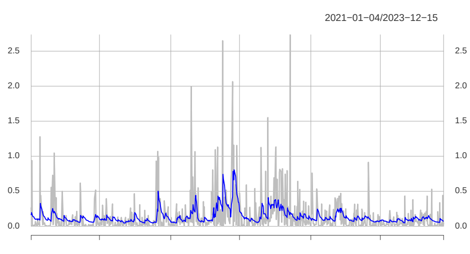
(d) New York



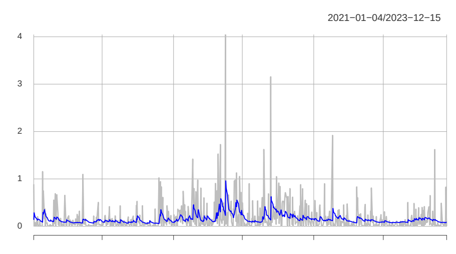
(e) Pennsylvania



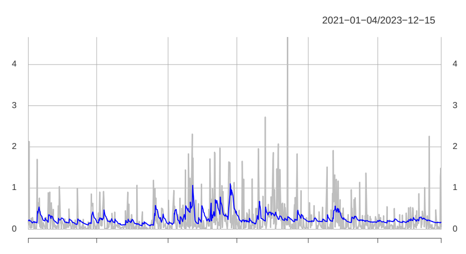
(f) Illinois



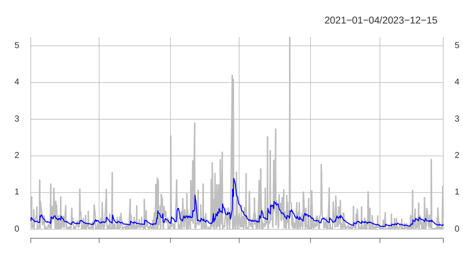
(g) Ohio



(h) Georgia



(i) North Carolina



(j) Michigan

3.3 Cluster analysis and validity indices

To gain further insight into the analysis, we find clusters of homogeneous regions among the US states according to volatility predictions. The algorithm is used here to construct a hierarchy of clusters, starting from the volatility forecasts of the EN-Comb and GJR models. A detailed description of the hierarchical clustering algorithm is beyond the scope of this work (see, for example, Murtagh and Contreras, 2012, 2017, for literature reviews, among others). Hereafter we make use the dissimilarity matrix calculated using the DTW metric (as described in Franses and Wiemann, 2020, among others) in the initial step. Subsequently, clusters are merged based on the Ward's linkage criterion, whereby pairs of clusters are aggregated to minimize the total within-cluster variance.[6]

In the literature, several indices have been proposed for measuring the validity of each partition and calculating the optimal number of clustering. In this work, we use four validity indices, briefly described as follows:

Silhouette score (Rousseeuw, 1987). It is in interval $[-1;1]$ and represents a measure of how well the observation fits into its assigned cluster. A higher value indicates a better quality of clustering.

Dunn index (Dunn, 1974). It measures the compactness (intra-cluster distance) and separation (inter-cluster distance) of clusters. A higher value indicates a better quality of clustering.

Davies-Bouldin index (Davies and Bouldin, 1979). It is based on both the intra-cluster and inter-cluster distances, and measures the average similarity between each cluster and its most similar cluster. A lower value indicates a better quality of clustering.

Ball-Hall index (Ball and Hall, 1965). It is based on cluster mean dispersion, defined as the mean of the squared distances of the cluster's points with respect to their barycenter. The Ball-Hall index is the mean of all the clusters mean dispersions. A lower value indicates a better quality of clustering.

The cluster analysis reveals distinct patterns in the volatility forecasts produced by the EN-Comb and GJR models. Table 6 presents the validity indices for different cluster solutions. Notably, the EN-Comb model, incorporating EPU information, consistently produces clusters with higher validity across all indices compared to the baseline GJR model. The EN-Comb model achieves the highest Silhouette score, Dunn index, and lowest Davies-Bouldin and Ball-Hall indices across all cluster solutions, suggesting that it effectively captures the underlying structure of volatility patterns better than the baseline. This superior performance is attributed to the EN-Comb model's ability to integrate state-specific EPU data, allowing it to discern nuanced regional variations in volatility responses to policy-related uncertainty. Specifically, the EN-Comb model identifies distinct clusters of states with similar volatility profiles, potentially linked to shared economic structures or geographical proximity.

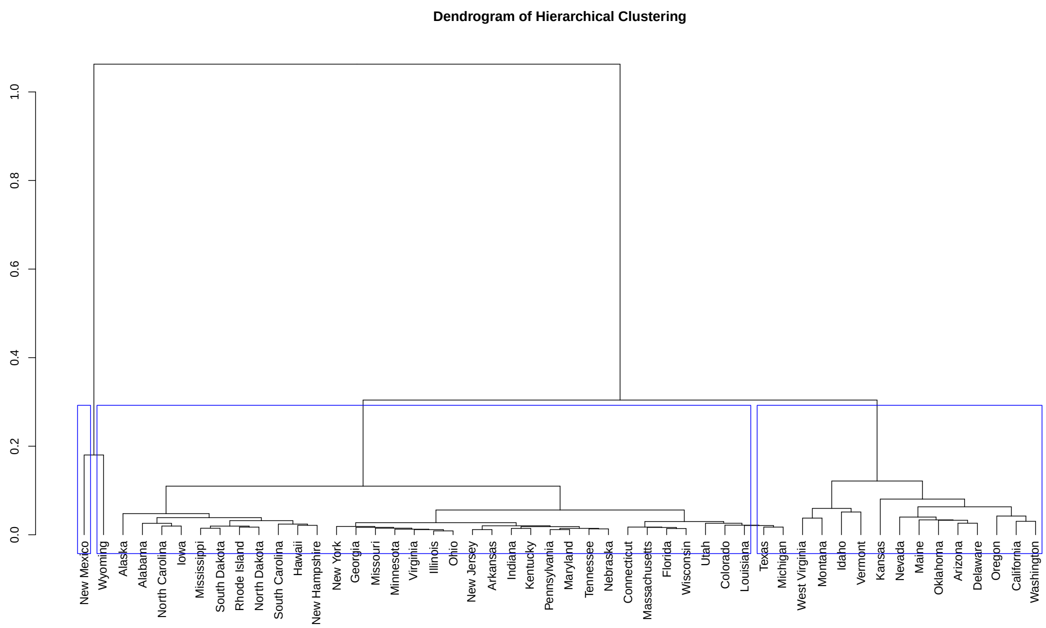
Table 6: Validity indices for clusters

Number of clusters	3		4		5	
	GJR	EN-Comb	GJR	EN-Comb	GJR	EN-Comb
Silhouette score (H)	0.482	0.487	0.464	0.470	0.403	0.403
Dunn index(H)	0.086	0.131	0.223	0.244	0.236	0.244
Davies-Bouldin index (L)	1.011	0.805	0.663	0.650	0.900	0.940
Ball-Hall index (L)	5.984	0.895	0.491	0.403	0.566	0.466

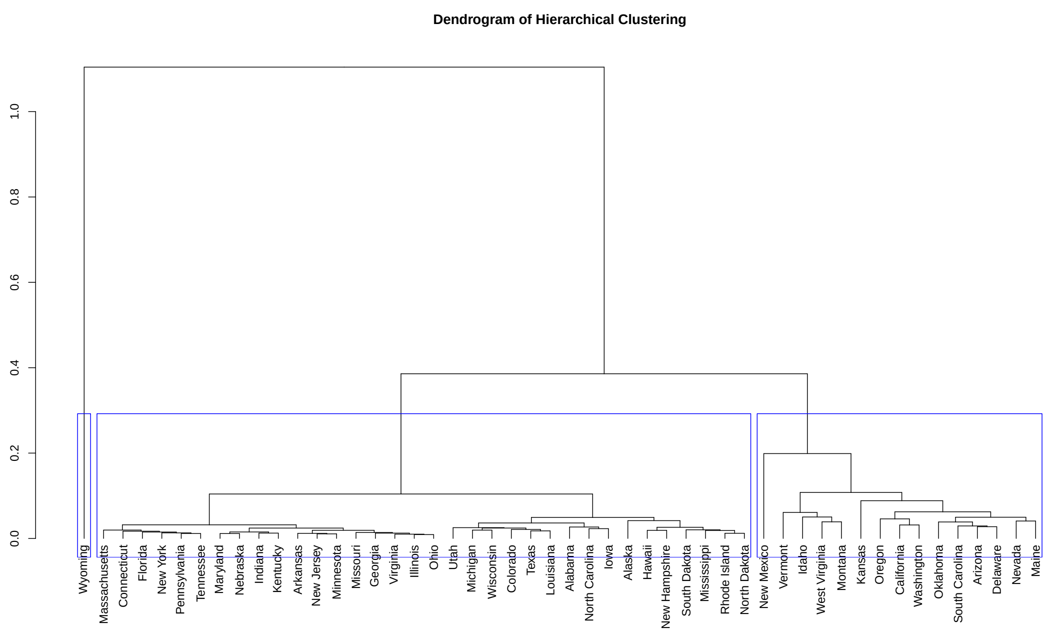
Notes: The table reports four internal validity indices for the clusters. (H) indicates that higher values represent better validity, while (L) indicates that lower values are better. Sample period: January 2021 to December 2023 (750 observations).

[6] The cluster analysis is performed using the R package *dtwclust* of Sardá-Espinosa (2019).

Figure 4: Hierarchical Cluster Dendrograms.

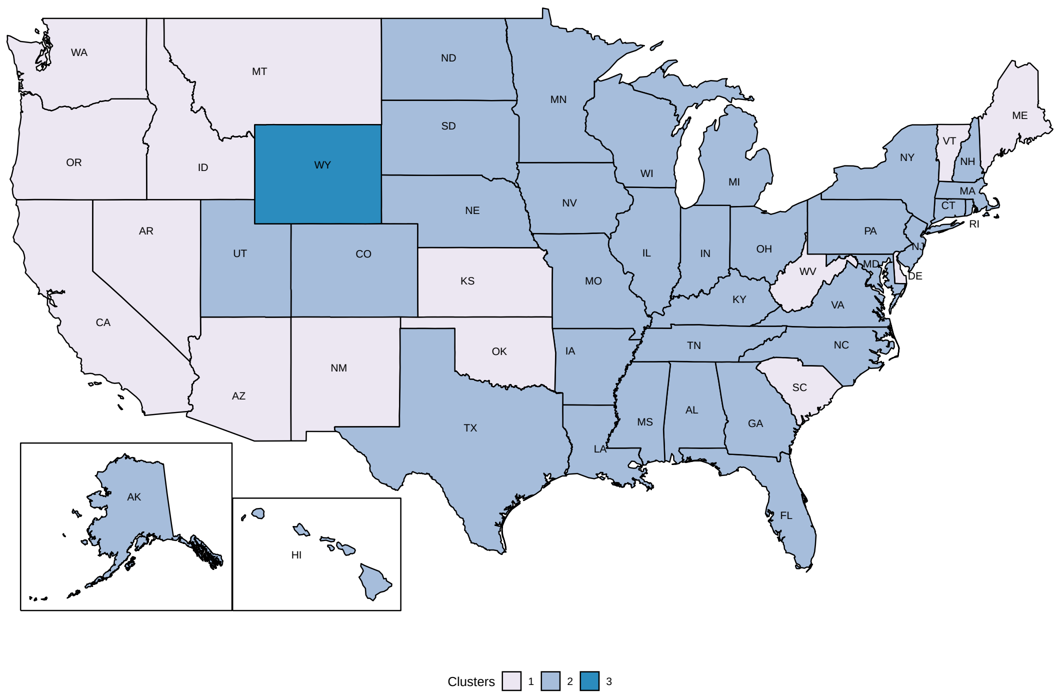


(a) GJR



(b) EN-Comb

Figure 5: Clusters of states for EN-Comb.



4 Conclusion

In this study we have documented the informational effectiveness from incorporating local and global EPU indices into the GARCH-MIDAS framework for forecasting US state-level equity returns volatility. By leveraging the Elastic Net combination approach, we find that our model generally outperforms both the benchmark GJR model and individual GARCH-MIDAS models using single EPU predictors. This underscores the value of integrating multiple sources of EPU information when forecasting volatility.

The decision to forecast state-level daily equity returns volatility at a mixed frequency using economic policy uncertainty indices combines the needs to provide both more accurate measures of volatility (Ghysels et al., 2019), and terms of reference for timely portfolio decisions and for risk management (Ghysels and Valkanov, 2012): as a matter of fact, more robust risk models and stress-testing scenarios enable risk managers to face potential economic downturns. By the same token, policymakers are able to gauge the impact of policy decisions on different regions and, possibly, tailor their interventions accordingly.

While our analysis focuses on the US context, the methodology and findings of this study can be extended to other regions with heterogeneous stock markets where EPU data is available. Future research could investigate the impact of EPU on stock market volatility in European countries, examining the distinct reactions across diverse economic and political landscapes, thus contributing to a global understanding of the EPU-volatility nexus. Additionally, exploring the potential of other mixed-frequency models or alternative combination strategies for incorporating EPU information could further improve volatility forecasting accuracy and enhance decision-making in financial markets.

Overall, this study provides convincing evidence for the predictive power of EPU indices in volatility forecasting at the state level. The GARCH-MIDAS framework, enhanced by the Elastic Net combination approach and cluster analysis, offers a comprehensive tool for understanding and navigating the complex relationship between economic policy uncertainty and stock market volatility.

References

Amado, C., Silvennoinen, A., and Teräsvirta, T. (2019). Models with multiplicative decomposition of conditional variances and correlations. In Chevallier, J., Goutte, S., Guerreiro, D., Saglio, S., and Sanhaji, B., editors, *Financial Mathematics, Volatility and Covariance Modelling*, volume 2, pages 217–260. Routledge.

Amado, C. and Teräsvirta, T. (2008). Modelling conditional and unconditional heteroskedasticity with smoothly time-varying structure. Technical Report 8, CREATES Research Paper.

Amendola, A., Braione, M., Candila, V., and Storti, G. (2020). A model confidence set approach to the combination of multivariate volatility forecasts. *International Journal of Forecasting*, 36(3):873–891.

Amendola, A., Candila, V., and Gallo, G. M. (2019). On the asymmetric impact of macro–variables on volatility. *Economic Modelling*, 76:135–152.

Amendola, A., Candila, V., and Gallo, G. M. (2021). Choosing the frequency of volatility components within the Double Asymmetric GARCH–MIDAS–X model. *Econometrics and Statistics*, 20:12–28.

Amendola, A. and Storti, G. (2008). A GMM procedure for combining volatility forecasts. *Computational Statistics & Data Analysis*, 52(6):3047–3060.

Amendola, A. and Storti, G. (2015). Model uncertainty and forecast combination in high-dimensional multivariate volatility prediction. *Journal of Forecasting*, 34(2):83–91.

Andersen, T. G. and Benzoni, L. (2009). Realized volatility. In Andersen, T. G., Davis, R. A., Kreiss, J. P., and Mikosch, T., editors, *Handbook of Financial Time Series*. Springer Verlag.

Asgharian, H., Hou, A. J., and Javed, F. (2013). The importance of the macroeconomic variables in forecasting stock return variance: A GARCH-MIDAS approach. *Journal of Forecasting*, 32(7):600–612.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Baker, S. R., Davis, S. J., and Levy, J. A. (2022). State-level economic policy uncertainty. *Journal of Monetary Economics*, 132(C):81–99.

Balcilar, M., Gupta, R., and Pierdzioch, C. (2016). Does uncertainty move the gold price? New evidence from a nonparametric causality-in-quantiles test. *Resources Policy*, 49(C):74–80.

Ball, G. H. and Hall, D. J. (1965). *ISODATA, a novel method of data analysis and pattern classification*, volume 699616. Stanford research institute Menlo Park, CA.

Bates, J. and Granger, C. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 30:451–468.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327.

Bollerslev, T. and Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11:143–172.

Carnero, M. A., Peña, D., and Ruiz, E. (2007). Effects of outliers on the identification and estimation of GARCH models. *Journal of Time Series Analysis*, 28(4):471–497.

Chaney, T., Sraer, D., and Thesmar, D. (2012). The collateral channel: How real estate shocks affect corporate investment. *American Economic Review*, 102(6):2381–2409.

Chaudhuri, T. D. and Ghosh, I. (2015). Using Clustering Method to Understand Indian Stock Market Volatility. *Communications on Applied Electronics*, 2(6):35–44.

Cipollini, F., Gallo, G. M., and Otranto, E. (2021). Realized volatility forecasting: Robustness to measurement errors. *International Journal of Forecasting*, 37(1):44–57.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41:135–155.

Conrad, C. and Loch, K. (2015). Anticipating long-term stock market volatility. *Journal of Applied Econometrics*, 30(7):1090–1114.

Copeland, T. E. (1976). A model for asset trading under the assumption of sequential information arrival. *Journal of Finance*, 31(4):114–1168.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7:174–196.

Coval, J. D. and Moskowitz, T. J. (1999). Home bias at home: Local equity preference in domestic portfolios. *Journal of Finance*, 54(6):2045–2073.

Coval, J. D. and Moskowitz, T. J. (2001). The geography of investment: Informed trading and asset prices. *Journal of Political Economy*, 199(4):811–841.

Davies, D. L. and Bouldin, D. W. (1979). A cluster separation measure. *IEEE transactions on pattern analysis and machine intelligence*, (2):224–227.

Dunn, J. C. (1974). Well-separated clusters and optimal fuzzy partitions. *Journal of cybernetics*, 4(1):95–104.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50:987–1007.

Engle, R. F. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics*, 17:425–446.

Engle, R. F., Ghysels, E., and Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics*, 95(3):776–797.

Engle, R. F. and Rangel, J. G. (2008). The spline-GARCH model for low frequency volatility and its global macroeconomic causes. *Review of Financial Studies*, 21:1187–1222.

Fang, T., Lee, T.-H., and Su, Z. (2020). Predicting the long-term stock market volatility: A GARCH-MIDAS model with variable selection. *Journal of Empirical Finance*, 58:36–49.

Franses, P. H. and Wiemann, T. (2020). Intertemporal similarity of economic time series: An application of dynamic time warping. *Computational Economics*, 56:59–75.

Ghysels, E., Plazzi, A., Valkanov, R., Rubia, A., and Dossani, A. (2019). Direct versus iterated multiperiod volatility forecasts. *Annual Review of Financial Economics*, 11(1):173–195.

Ghysels, E. and Qian, H. (2019). Estimating MIDAS regressions via OLS with polynomial parameter profiling. *Econometrics and Statistics*, 9:1–16.

Ghysels, E. and Valkanov, R. (2012). *Forecasting Volatility with MIDAS*, chapter 16, pages 383–401. John Wiley & Sons, Ltd.

Glosten, L. R., Jagannanthan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801.

Gong, X., Zhang, W., Xu, W., and Li, Z. (2022). Uncertainty index and stock volatility prediction: evidence from international markets. *Physica A: Statistical Mechanics and Its Applications*, 8(1):57.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The Model Confidence Set. *Econometrica*, 79(2):453–497.

Hoaglin, D. C. and Iglewicz, B. (1987). Fine-tuning some resistant rules for outlier labeling. *Journal of the American statistical Association*, 82(400):1147–1149.

Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67.

Korniotis, G. M. and Kumar, A. (2013). State-level business cycles and local return predictability. *Journal of Finance*, 68(3):1037–1096.

Li, D., Zhang, L., and Li, L. (2023). Forecasting stock volatility with economic policy uncertainty: A smooth transition GARCH-MIDAS model. *International Review of Financial Analysis*, 88(C):102708.

Liu, L. and Zhang, T. (2015). Economic policy uncertainty and stock market volatility. *Finance Research Letters*, 15(C):99–105.

Liu, Z., Ye, Y., Ma, F., and Liu, J. (2017). Can economic policy uncertainty help to forecast the volatility: A multifractal perspective. *Physica A: Statistical Mechanics and Its Applications*, 482(C):181–188.

Luo, Z., Zhang, L., Liu, N., and Wu, Y. (2023). Time series clustering of COVID-19 pandemic-related data. *Data Science and Management*, 6(2):79–87.

Murtagh, F. and Contreras, P. (2012). Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(1):86–97.

Murtagh, F. and Contreras, P. (2017). Algorithms for hierarchical clustering: an overview, II. *Data Mining and Knowledge Discovery*, 7(6):e1219.

Murtagh, F. and Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classification*, 31:274–295.

Nanda, S., Mahanty, B., and Tiwari, M. (2010). Clustering Indian stock market data for portfolio management. *Expert Systems with Applications*, 37(12):8793–8798.

Pástor, L. and Veronesi, P. (2012). Uncertainty about government policy and stock prices. *Journal of Finance*, 67(4):1219–1264.

Pástor, L. and Veronesi, P. (2013). Political uncertainty and risk premia. *Journal of Financial Economics*, 110(3):520–545.

Pirinsky, C. and Wang, Q. (2006). Does corporate headquarters location matter for stock returns? *Journal of Finance*, 61(4):1991–2015.

Poon, S. and Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 91:478–539.

Rapach, D. E., Strauss, J. K., and Wohar, M. E. (2008). Forecasting stock return volatility in the presence of structural breaks. In Rapach, D. and Wohar, M., editors, *Forecasting in the Presence of Structural Breaks and Model Uncertainty*, volume 3, pages 381–416. Emerald, Bingley, United Kingdom.

Raza, S. A., Shah, N., and Shahbaz, M. (2018). Does economic policy uncertainty influence gold prices? Evidence from a nonparametric causality-in-quantiles approach. *Resources Policy*, 57(C):61–68.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65.

Sakoe, H. and Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing*, 26(1):43–49.

Salisu, A. A., Demirer, R., and Gupta, R. (2023). Policy uncertainty and stock market volatility revisited: the predictive role of signal quality. *Journal of Forecasting*, 42(8):2307–2321.

Salisu, A. A., Liao, W., Gupta, R., and Cepni, O. (2024a). Economic Conditions and Predictability of US Stock Returns Volatility: Local Factor versus National Factor in a GARCH-MIDAS Model. *forthcoming in Journal of Forecasting*.

Salisu, A. A., Ogbonna, A. E., Gupta, R., and Bouri, E. (2024b). Energy-related uncertainty and international stock market volatility. *The Quarterly Review of Economics and Finance*, 95(C):280–293.

Sardá-Espinosa, A. (2019). Time-Series Clustering in R Using the dtwclust Package. *The R Journal*, 11(1):22.

Segnon, M., Gupta, R., and Wilfling, B. (2024). Forecasting stock market volatility with regime-switching GARCH-MIDAS: The role of geopolitical risks. *International Journal of Forecasting*, 40(1):29–43.

Shiller, R. J. (1981a). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71(3):421–436.

Shiller, R. J. (1981b). The use of volatility measures in assessing market efficiency. *Journal of Finance*, 36(2):291–304.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288.

Timmermann, A. (2006). Forecast combinations. volume 1 of *Handbook of Economic Forecasting*, pages 135–196. Elsevier, Amsterdam, Netherlands.

Wang, J., Ma, F., Bouri, E., and Guo, Y. (2023a). Which factors drive bitcoin volatility: Macroeconomic, technical, or both? *Journal of Forecasting*, 42(4):970–988.

Wang, X., Hyndman, R. J., Li, F., and Kang, Y. (2023b). Forecast combinations: an over 50-year review. *International Journal of Forecasting*, 39(4):1518–1547.

Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244.

Zhang, Y., Wei, Y., Zhang, Y., and Jin, D. (2019). Forecasting oil price volatility: Forecast combination versus shrinkage method. *Energy Economics*, 80:423–433.

Zou, H. (2006). The Adaptive Lasso and its Oracle Properties. *Journal of the American Statistical Association*, 101(476):1418–1429.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2):301–320.

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.unica.it

24/13 *Ester Rita Lucia Sanna, Marta Meleddu,* "Integrating justice and ecological economics: a theoretical framework and indicator toolkit for analysing conflict in protected areas"
24/12 *Federico Aresu, Emanuela Marrocu, Raffaele Paci,* "EU funds and TFP growth: how the impact changed over time and space"
24/11 *Ivan Etzo, Raffaele Paci, Cristian Usala,* "Brain gain vs. brain drain. The effects of universities' mobile students on territorial inequalities"
24/10 *Giommaria Spano,* "Financial frictions, common ownership and firms' market power in a general equilibrium model"
24/09 *Roberto Saba, Anna Pireddu,* "Energy and Environmental Synergy: Cooperative Strategies for the Euro-Mediterranean Transition"
24/08 *Maria Grazia Pittau, Francesco Politano, Roberto Zelli,* "Educational divide between voters: A nationwide trend?"
24/07 *Federico Atzori, Vittorio Pelligra,* "Framed norms. The effect of choice-belief information on tax compliance"
24/06 *Simone Nobili,* "Concentration, Market Power and International Tax Competition"
24/05 *Maria Grazia Pittau, Pier Luigi Conti, Roberto Zelli,* "Inference for deprivation profiles in a binary setting"
24/04 *Edoardo Otranto,* "A Vector Multiplicative Error Model with Spillover Effects and Co-movements"
24/03 *Luca Serafini, Emanuela Marrocu, Raffaele Paci,* "Smart Strategies, Smarter Performance: the Impact of S3 and Industry 4.0 on Firms' Outcomes"
24/02 *Hugues Champeaux, Elsa Gautrain, Karine Marazyan,* "Men's premarital migration and marriage payments: Evidence from Indonesia"
24/01 *Pablo Álvarez-Aragón, Hugues Champeaux,* "Measuring Norms and Enumerator Effects: Survey Method Matters"
23/15 *Fabio Cerina, Xavier Raurich,* "Saving Behaviour and the Intergenerational Allocation of Leisure Time"
23/14 *Leonardo Vargiu, Bianca Biagi, Maria Giovanna Brandano, Paolo Postiglione,* "Research Infrastructures and Regional Growth: the case of Europe"
23/13 *Alberto Tidu,* "Dissecting inequality: conceptual problems, trends and drivers"
23/12 *Luciano Mauro, Francesco Pigliaru,* "Italy's National Recovery and Resilient Plan: Will it Narrow the North-South Productivity Gap?"
23/11 *Fabrizio Antolini, Samuele Cesarini, Giorgio Garau,* "The economic impact of the tourism sector on the overall Italian economy: An Input-Output Approach"
23/10 *Giorgio Garau, Andrea Karim El Meligi,* "The Impact of the Pandemic and War on Surplus Redistribution Mechanisms: A Sectoral Analysis of France and Italy"
23/09 *Maria Giovanna Brandano, Alessandra Faggian, Adriana C. Pinate,* "The impact of COVID-19 on the tourism sector in Italy: a regional spatial perspective"
23/08 *Fabrizio Cipollini, Giampiero M. Gallo, Alessandro Palandri,* "Modeling and evaluating conditional quantile dynamics in VaR forecasts"
23/07 *Ugo M. Gragnolati, Luigi Moretti, Roberto Ricciuti,* "Early railways and industrial development: Local evidence from Sardinia in 1871–1911"
23/06 *Giampiero M. Gallo, Demetrio Lacava, Edoardo Otranto* "Volatility jumps and the classification of monetary policy announcements"

www.crenos.unica.it

ISBN 9788868515416