



Identifying market regimes via asset class correlations

By Editor October 28, 2022 👁 4,978

A recent paper suggests identifying financial market regimes through the correlations of asset class returns. The basic idea is to calculate correlation matrixes for sliding time windows and then estimate pairwise similarities. This gives a matrix of similarity across time. One can then perform principal component analysis on this similarity matrix and extract the "axes" of greatest relevance. Subsequently, one can cluster the dates in the new reduced space, for example by a K-means method, and choose an optimal number of clusters. These clusters would be market regimes. Empirical analyses of financial markets over the last 20-100 years identify 6-7 market regimes.

Cucuringu Mihai and Deborah Miori (2022), "Returns-Driven Macro Regimes and Characteristic Lead-Lag Behaviour between Asset Classes."

The below post consists of quotes from the paper. It ties in with this site's [summary on quantitative methods for macro information efficiency](#), particularly the section on unsupervised learning.

The basic idea

"The broader dynamics of financial markets can significantly vary across time. These can define periods of different macroeconomic regimes, which are clustered moments of persistent market conditions that can be characterised by external macroeconomic trends. The importance of regime identification mainly lies in its implications and impact on asset allocation and portfolio construction."

"We pursue a detailed investigation...of return correlations and causalities for securities belonging to multiple asset classes...fuelled by its strong potential to shed further light on macro regimes identification and characterisation...We define data-driven macroeconomic regimes by clustering the relative performance of indices belonging to different asset classes [across time]. We then investigate lead-lag relationships within the regimes identified."

How to identify market regimes based on correlation

"Our first aim is to identify a reliable division of time into financial regimes from a data-driven point of view, partly...Each regime is then characterised by summary statistics and relationships between the returns of different asset classes."

"Market similarity in time and regimes [is identified through] correlation of asset returns. For each sliding window of time, we compute Pearson linear correlation, Spearman monotonic correlation, and Kendall rank correlation between the time series of indices' returns...After obtaining a set of correlation matrices for different points in time, we compute the related pairwise similarities. We consider two relevant measures:

1. **Cophenetic correlation**:...We transform each correlation matrix into a distance matrix...Then, we pursue hierarchical clustering via the linkage algorithm with the average method...Finally, the cophenetic correlation is a measure of the correlation between the distances of points in the feature space and on the dendrogram generated by the clustering.
2. **Metacorrelation**:...We flatten the matrices of assets' correlations and, for each pair, compute Pearson's correlation."

"Next, we perform Principal Component Analysis (PCA) on the resultant matrix of similarity between points in time. This is done to extract the axes that explain the greatest variance between the data and project the latter on them. We cluster dates in this new space via KMeans++ to identify different financial regimes. The optimal number of groups/regimes is chosen following the elbow method, i.e. looking for the rough optimal point with lowest inertia and lowest number of clusters on the related plot. Inertia measures the distance between each data point and its centroid, squares it, and sums these results across each one cluster."

"To increase confidence in the identified regimes, we compare them with the results [of a] network study of the evolution of correlations...We start from our set of correlation matrices between assets' returns at different points in time and retain only entries with magnitude above a threshold. From each matrix, we build the related signed network...The set of nodes are indices and edges...We save the giant components and cluster each final graph via...a spectral algorithm that aims at having the maximum number of positive edges within clusters while having the highest number of negative edges between clusters. To choose the number of communities to look for, we compute the signed modularity [a measure of the structure of networks or graphs which measures the strength of division of a network into modules]."

"Once regimes have been identified, we proceed by studying their peculiar features. We compute the average correlation matrix for each regime by considering the points in time assigned to it. Then, we also compute the related average mean returns, average standard deviations of returns and annualised Sharpe Ratios of the underlying asset classes."

How to use identified regimes for trading

"The second step of our analysis investigates lead-lag relationships between clusters of assets, for each uncovered regime. The results are then tested for the profitability of an investment on the lagging assets."

"After having identified our macro regimes, we can investigate whether any latent significant lead-lag relationships could be extracted. Thus, we fragment our time series of indices' returns into the related regimes. For each regime, we compute the Granger causality [predictive power of one

time series with respect to another] between each pair of indices for [various] lags trading days and retain results at the 0.05 significance level."

"Each resultant causality matrix is used to build a directed network for the related regime, from which we can identify leading and lagging groups of indices via the Hermitian clustering algorithm."

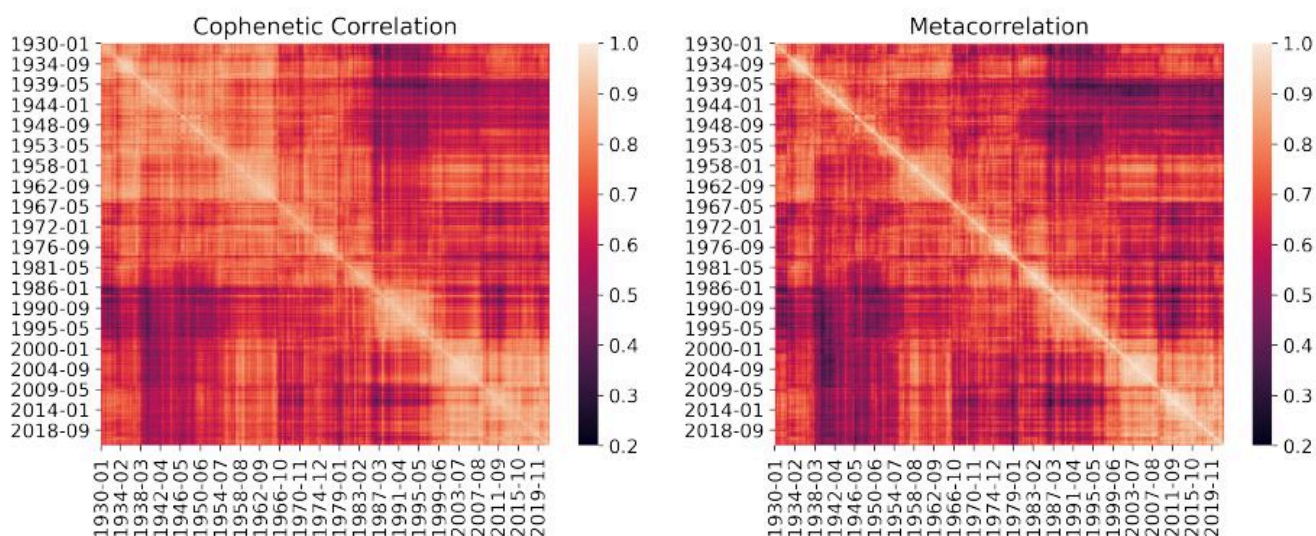
"As a final step, we compare the performance of an investment based on lead-lag clusters specific to each regime against a vanilla benchmark that bets uniformly on all assets...We consider the most leading cluster in each regime, compute the average return of the related subset of indices over the past days, and use its sign as a signal to buy or sell assets uniformly in the most lagging cluster. "

An empirical illustration

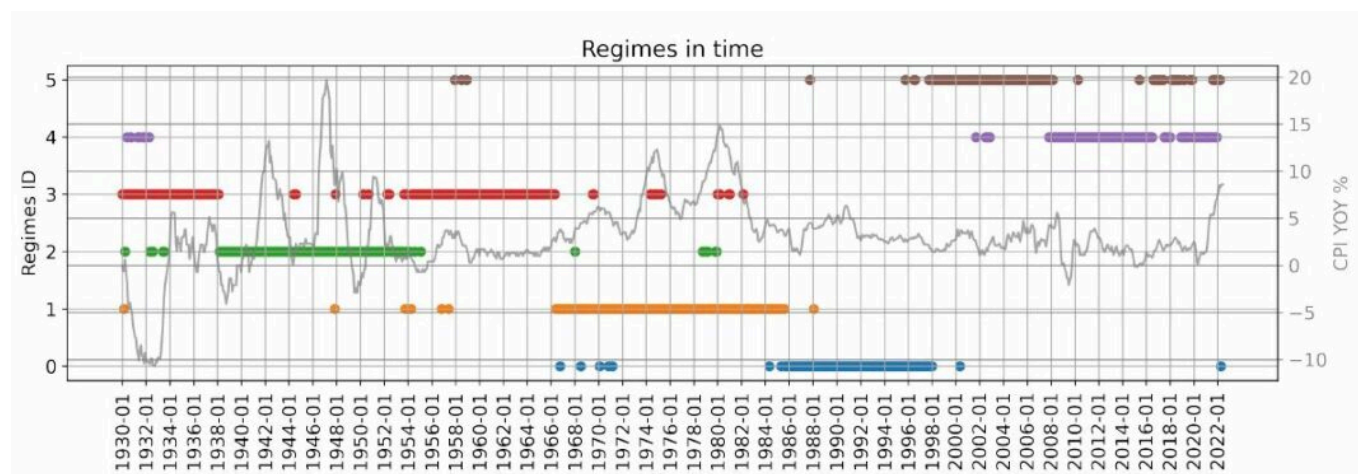
"The first data set considered is provided by Fidelity Investments...an internal set of 33 indices belonging to different asset classes, whose majority of levels have been reconstructed at the last day of each month, starting from January 1921...Indices relate to equity, commodity, fixed income and a proxy for cash."

"[The second data set consists of] daily levels of a broad set of indices belonging to different asset classes from Bloomberg...which belong to the class of commodities, currencies, equities, bond spreads, volatilities and interest rates...The data start on 30 September 2005, and end on 1 July 2022."

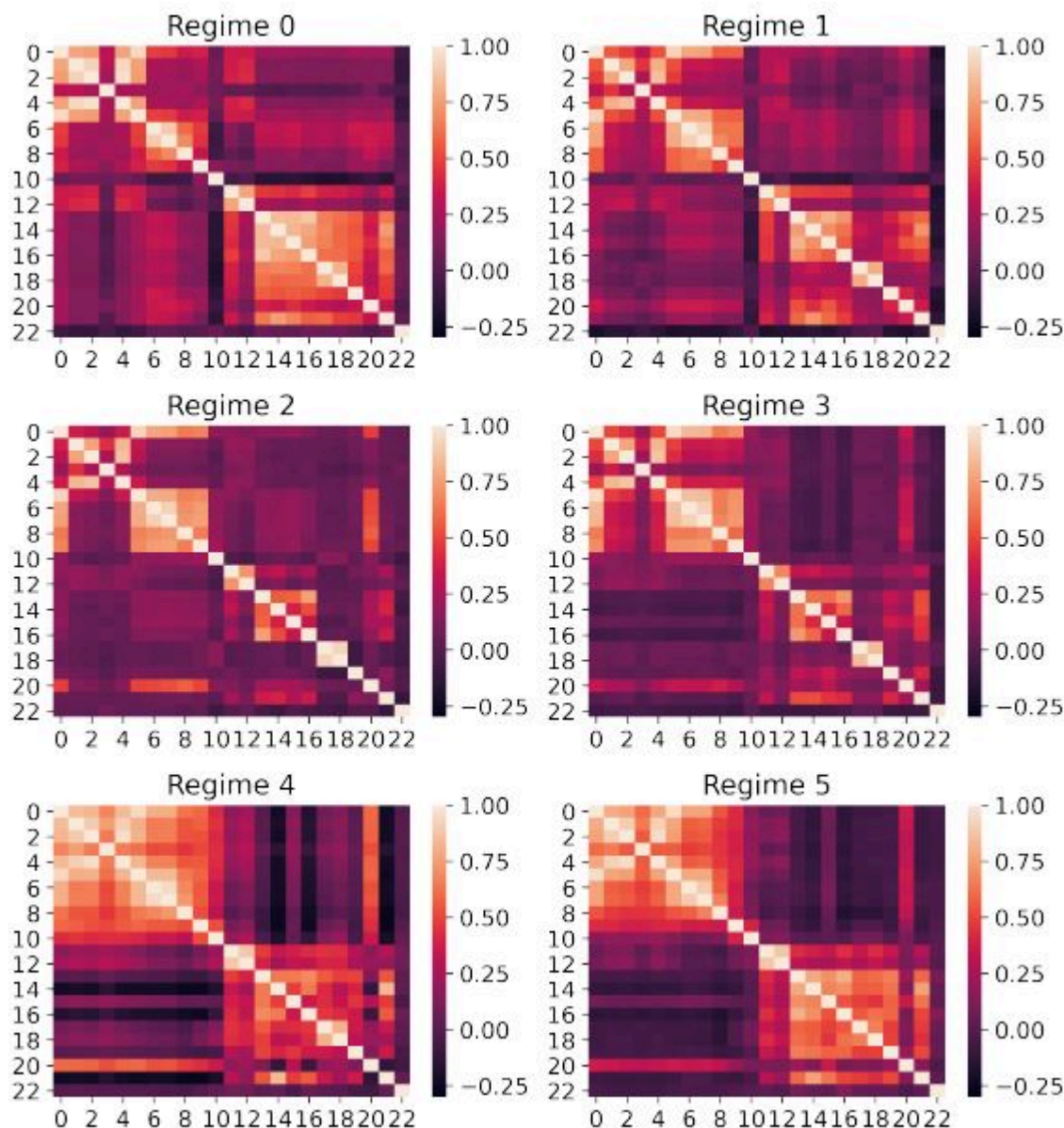
"The time-similarity matrices [of the figure below for the Fidelity series] show a clear block structure that already hints to a related division of periods into regimes. We perform PCA on these matrices and find that three dimensions explain 92% of the variance within the data for the cophenetic case, while four dimensions account for 91% of it for metacorrelations."



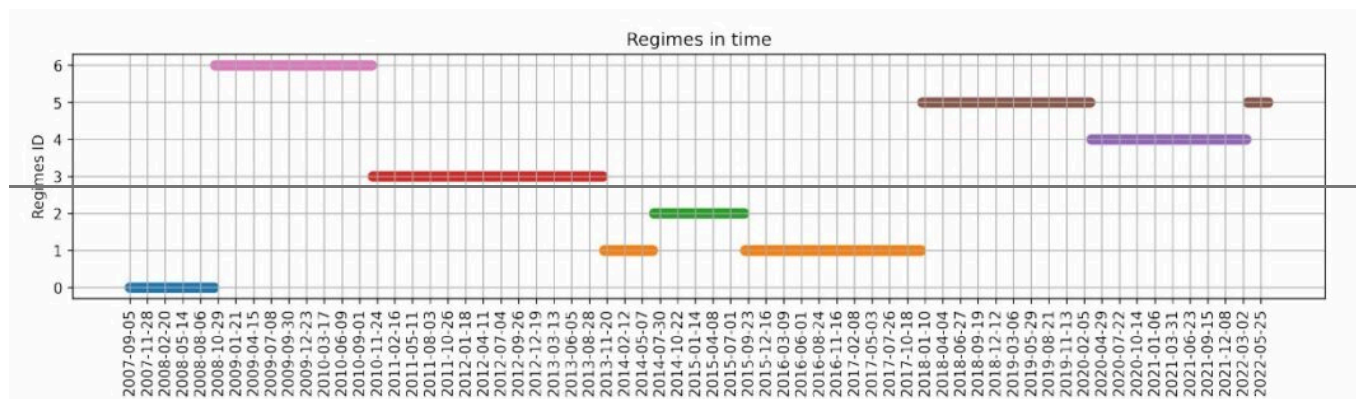
"Next, we cluster points projected onto each new space via KMeans++. The best number of regimes is chosen following the elbow method on an inertia plot, but the stability of results is checked against perturbations of the amount of desired clusters. We also confirm the overall stability of our results by varying initial random seed, time window length and number of dimensions kept in PCA. Our results suggest the existence of six regimes, shown in the figure below from the metacorrelation matrix. We also plot inflation as Consumer Price Index Year-Over-Year in percentages (CPI YOY %) to have a macroeconomic indicator for comparison."



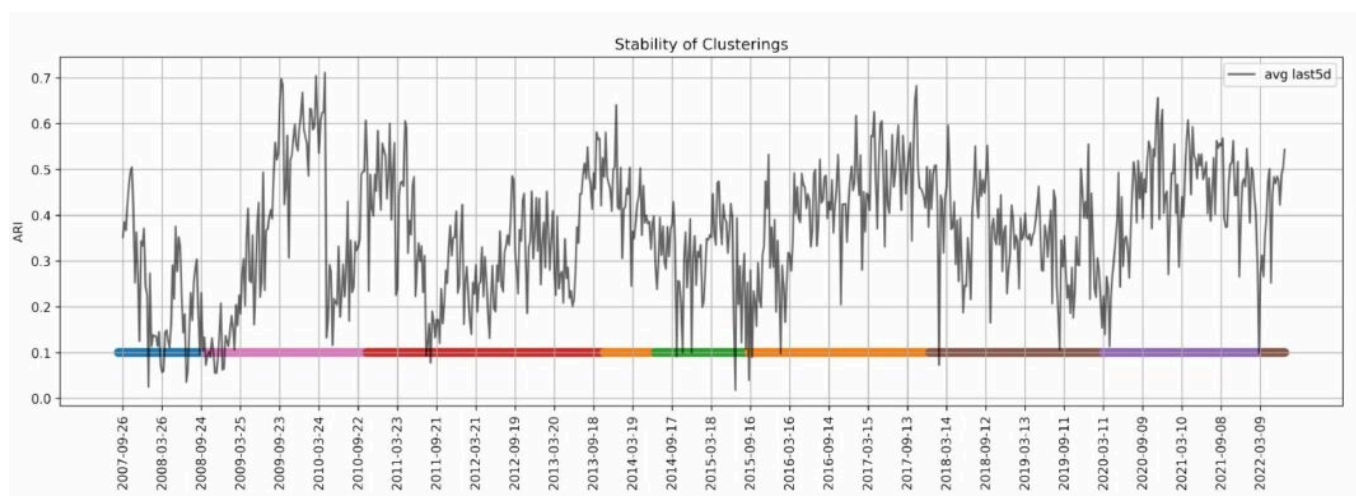
"Finally, we calculate the average correlation matrix between assets for each different regime. The results are shown [in the figure below]."



"We [also] compute correlations between Bloomberg indices' returns using a window of two years length that slides one week at every iteration. Motivated by similar considerations to the ones for the Fidelity data set, we focus on the extraction of regimes from correlations. The matrix of periods' similarities shows again a clear block structure...We do PCA and see that three dimensions describe ~ 90% of the variance of the data. Points in time are projected onto this new space and clustered via KMeans, where the optimal number of groups is again chosen by looking at the inertia loss function. We find an optimal discretisation of time into seven regimes, which are [in the figure below]."



“Network-based identification of regimes: We increase the confidence in our regimes by separately studying the structural evolution in time of the network of returns’ correlations between indices. We build one network per week following the adopted sliding window. Nodes are the Bloomberg indices and weighted edges are added from the related correlations if their magnitude is above the threshold. Then, the giant component of each network is kept. A higher number of survival links is an indicator of increasing correlations between assets and points towards periods of market distress...We proceed by clustering nodes of each signed network...The similarity between communities extracted at consequent points in time is computed by averaging the ARI [Adjusted Rand Index] value between the current clustering and each clustering for the past four weeks. The result is plotted in the figure below.”



“We compare the average return of a portfolio constructed by considering the most leading and lagging clusters versus a plain investment on all indices [using in-sample information]...In each regime, we have positive return (reported in percentage) that also outperform the benchmark... While this is not a real investment strategy, the framework shows the importance of recognising both regimes and their inner evolving leaders and laggards.”