Detection of Regime Changes in the US Financial Market

Benjamin Li, Charlotte Peterson, Andrew Park, Edmond Li, James Oh, Willis Ng, Aaron Noh, David Kim, Marco Giordano, Paul Lee, Roy Kang, Raphael Douady, Tracy Strickel

Keywords: Breakpoint Detection, Regime Switching Model, Econometrics, Machine Learning, Quantitative Finance

Abstract

The aim of this work is to detect regime change in financial markets. Regime change refers to a substantial shift in the behavior of financial markets. We track change in market regime caused by surges in volatility by computing the Mahalanobis distance between a vector of daily returns and a reference covariance matrix. We also track changes in market regime caused by either surging or calming of volatility, or by trend reversal, by computing the Hellinger distance between a distribution representing daily asset returns and a reference distribution representing previous market behavior. When the resultant Hellinger distance exceeds a threshold, we say that the regime has changed. We note that while market volatility surges are easy to detect, because they occur brutally, volatility drops are much slower and occur in a very smooth way, making them much more difficult to identify precisely. Using this method, we are able to detect fourteen distinct market regimes between 2012 and 2023.

Introduction: Regime Change and Breakpoints

A financial market regime refers to an era of persistent market conditions. Regime switching models effectively capture the tendency of financial markets to abruptly change behavior, and the subsequent persistence of these changes. These models, pinpointing different regulatory, policy, and secular periods, identify distinct market states through econometric techniques (1).

They often reveal distinct periods that align with changes in regulations, policies, and other significant shifts over time. For instance, between 1979 and 1982, there was a notable shift in interest rate patterns coinciding with the Federal Reserve's adoption of a new approach to targeting monetary aggregates. Similarly, different interest rate patterns can be linked to specific tenures of different Federal Reserve Chairs. These regime switches also hold substantial implications for investors' optimal portfolio decisions (8).

Data

US financial market data (daily) was collected from Yahoo Finance from 1/28/2011 to 8/11/2023. We selected a universe of fifteen ETFs tracking the SP500 index, as well as fourteen industry indices: financial, utilities, industrial, technology, healthcare, consumer discretionary, consumer staples, materials, oil and gas, energy, metals and mining, retail, telecoms. For each ETF, we compute its trend, represented as the exponential moving average of its returns (with a decay factor of 0.99, i.e. a half-life a bit over 3 months), and the variance with the same exponential decay. Then, for each pair of ETFs, their covariance, also computed with the same exponential decay. At any given date, the distribution of returns is assumed to be a multivariate Gaussian distribution centered at the trend of the ETFs and with a variance-covariance matrix, both computed on that date.

Let $R_i(t)$ be the return of ETF X_i at date t and $\alpha = 0.99$ be the decay factor. We first compute the exponentially weighted moving average:

$$\mu_i(255) = \frac{1}{255}\sum_{i=1}^{255} R_i(t) \tag{1}$$

And for $t > 255$

$$\mu_i(t) = \alpha\mu_i(t-1) + (1-\alpha)R_i(t) \tag{2}$$

For any pair of ETFs (X_i, X_j) with possibly $i = j$, we also compute the uncentered covariance:

$$\gamma_{ij}(255) = \frac{1}{255}\sum_{i=1}^{255} R_i(t)R_j(t) \tag{3}$$

And for $t > 255$

$$\gamma_{ij}(t) = \alpha\gamma_{ij}(t-1) + (1-\alpha)R_i(t)R_j(t) \tag{4}$$

And finally, the centered covariance (and variance in the case $i = j$):

$$\sigma_{ij}(t) = \gamma_{ij}(t) - \mu_i(t)\mu_j(t) \tag{5}$$

Malahanobis Metric

The Mahalanobis metric (2), (3) measures the distance between a point $\vec{x} = (x_1, \ldots, x_{15})$ and a Gaussian distribution $Q = N(\vec{\mu}, \Sigma)$ on $\mathbb{R}^{15}$ with mean $\vec{\mu} = (\mu_1, \ldots, \mu_{15})$ and covariance matrix $\Sigma = \left(\sigma_{ij}\right)_{1 \leq i,j \leq 15}$:

$$d_M(\vec{x}, Q) = \sqrt{(\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \overrightarrow{\mu})} \tag{6}$$

In this case, the point $\vec{x}$ is the vector of returns $\overrightarrow{R_t}$ for all assets on day t, while Q is a distribution of returns chosen to be representative of recent market activity. The Mahalanobis metric corresponds to the number of standard deviations the vector $\vec{x}$ is away from the center of the distribution $\vec{\mu}$ (3).

We use the Mahalanobis metric between the vector of 15 returns $\vec{R}(t)$ for date t, and the distribution Q at the previous date, using the trend $\vec{\mu}(t-1)$ and covariance matrix $\Sigma(t-1)$, so as to capture the surprise effect of the current return with respect to the previous day statistics. When the Mahalanobis metric spikes, we consider this indicative of a surge of volatility and, potentially, a regime change.

If the returns were truly Gaussian-distributed, the square of the Mahalanobis metric d_M^2 would follow a χ-squared distribution with 15 degrees of freedom. Its expectation would be 15. In practice, due to the fat-tailed distribution of returns, in our sample, it is equal to 17.5 and that of d_M itself is 3.92.

We chose to consider that a market shift with $d_M > 10$ (about 2.5 times the average value, an event which, in our sample, occurs 0.6% of the time) represents a large enough shock to trigger a regime change.

It often occurs that, following a large move in terms of Mahalanobis metric, the market becomes agitated and, before the covariance matrix $\Sigma(t)$ fully factors in this agitation, the Mahalanobis metric of market moves during the next coming days remains large. In this case, we only consider the first surge of the metric to be indicative of a breakpoint.

Hellinger distance

The Mahalanobis metric captures deviations from previous market behavior and can thus detect regime change resulting from surges in market volatility (4). In theory, it could also detect a sudden trend reversal but, in practice, it is rare that such an event occurs without the volatility surging concomitantly. However, it cannot capture a regime change resulting from market volatility calming, because such an event would not be seen as exceptional. We attempt to detect this type of regime change using the Hellinger distance.

The Hellinger distance H between two distributions Q_1 and Q_2 with covariance matrices Σ_1 and Σ_2 and means $\vec{\mu}_1$ and $\vec{\mu}_2$ is given by (5) and (7):

$$H(P,Q) = 1 - \frac{|\Sigma_1|^{\frac{1}{4}}|\Sigma_2|^{\frac{1}{4}}}{|\bar{\Sigma}|^{\frac{1}{2}}} \exp\left(-\frac{1}{8}(\vec{\mu}_1 - \vec{\mu}_2)^T \bar{\Sigma}^{-1}(\vec{\mu}_1 - \vec{\mu}_2)\right) \quad (7)$$

where $\bar{\Sigma} = \frac{\Sigma_1+\Sigma_2}{2}$ and $|\Sigma|$ is the determinant of the matrix Σ.

The Hellinger distance shows how different the two distributions are (6). It is null when the two distributions are equal and its maximum value is 1, which is the case when the two distributions are foreign to each other, that is, when there an event of probability 1 for one distribution and 0 for the other (7). When the distance exceeds 0.5, we may consider that the distributions are more dissimilar than similar.

We measure the Hellinger distance between the distribution $Q = Q(t)$ at date *t* and that 20 trading days before (approximately one month), $Q' = Q(t-20)$, both represented by their trend and covariance matrix. Since the decay factor is $\alpha = 0.99$, we have:

$$Q = \alpha^{20}Q' + Q'' \quad (8)$$

Where the matrix Q'' depends on the returns in the last 20 days. Given that $\alpha^{20} \approx 0.8$, if the update Q'' is dissimilar enough from Q', then the Hellinger distance between Q and Q' should be at least $(1-0.8) \times 0.5 = 0.1$. Obviously, once the distance between $Q(t)$ and $Q(t-20)$ exceeds the threshold, this is likely to happen the next day. So, we only identify *t* as a breakpoint on dates where the threshold is passed, but is wasn't passed the previous day.

Since it often occurs that the Mahalanobis metric and the Hellinger distance surge more or less at the same time, in order to not identify regime changes that too narrowly spaced in time, we set a minimum time period of 60 days (approximately 3 months) between regime changes.

Results

Mahalanobis Surges

Table 1 shows the 8 dates at which the Mahalanobis metric exceeds the threshold 10, excluding a sequence of dates in early March 2020, following the first surge on 2/28/20, due to the COVID crisis (in particular 3/9/20 and 3/16/20 where it exceeded 20).

	1	2	3	4	5	6	7	8
Date	11/28/14	11/9/16	10/27/17	2/5/18	2/28/20	11/9/20	1/27/21	5/18/22
Mahalanobis	10.12	11.88	10.34	11.49	12.55	12.44	12.75	10.92

Table 1: Dates of Mahalanobis metric surges.

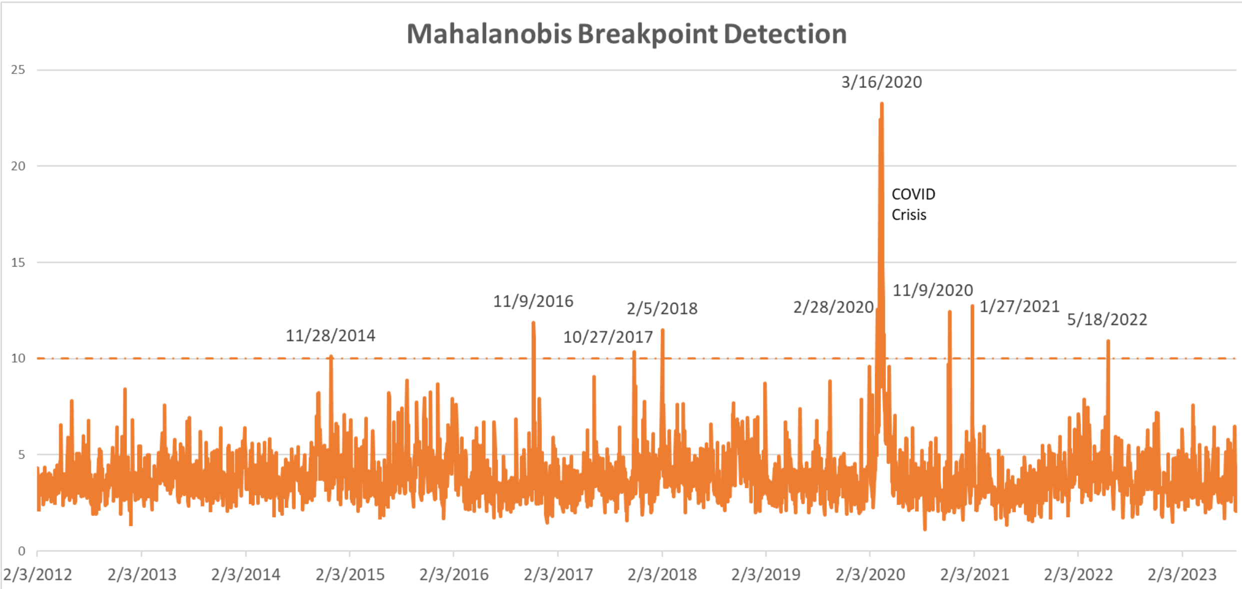

Figure 1: Mahalanobis metric of the SP500 index and its 14 industry indices

Hellinger Distance Surges

Table 2 shows the 9 dates at which the Hellinger distance with the distribution 20 days before passes above the threshold 0.1, excluding days that are less than 60 days after a previous passing of the threshold.

	1	2	3	4	5	6	7	8	9
Start Date	10/15/14	8/27/15	2/8/16	11/10/16	2/5/18	10/25/18	2/27/20	1/27/21	3/1/22
Peak Date	10/27/14	9/1/15	2/17/16	12/1/16	2/16/18	11/2/18	3/25/20	2/2/21	3/10/22
Hellinger	0.14	0.11	0.11	0.14	0.17	0.13	0.78	0.27	0.16

Table 3: Dates of Hellinger distance surges

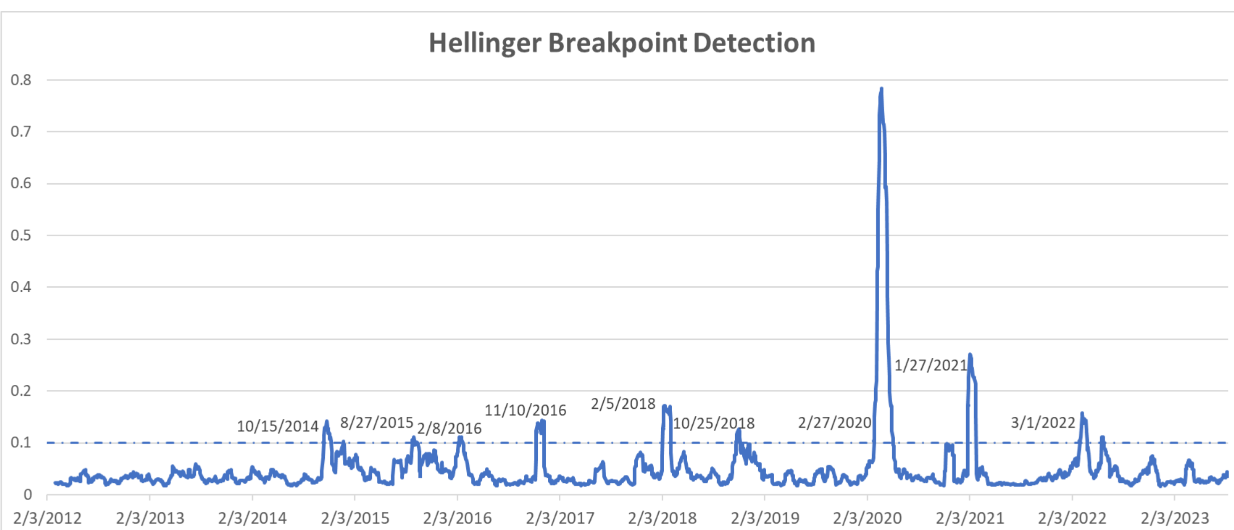

Figure 2: Hellinger distance of fitted Gaussian distributions 20 days apart

Synthesis of Mahalanobis and Hellinger breakpoints detection

Table 3 shows the regime start and end dates obtained using both the Mahalanobis and Hellinger metrics. We detected 10 breakpoints in the history of the SP500 index and its industry sub-indices during the period 2012-2023. Six of these dates were detected by the Hellinger criterion, while 3 others were detected by the Mahalanobis criterion, and one date was detected by both criteria. Most of the dates detected by one criterion were shortly followed by a date triggering the other criterion. This demonstrates internal consistency between the results of our two methods.

	1	2	3	4	5	6	7	8	9	10
Breakpoint	10/15/14	8/27/15	2/8/16	11/9/16	10/27/17	2/5/18	10/25/18	2/27/20	11/9/20	3/1/22
Criterion	Hellinger	Hellinger	Hellinger	Mahalanobis	Mahalanobis	Both	Hellinger	Hellinger	Mahalanobis	Hellinger

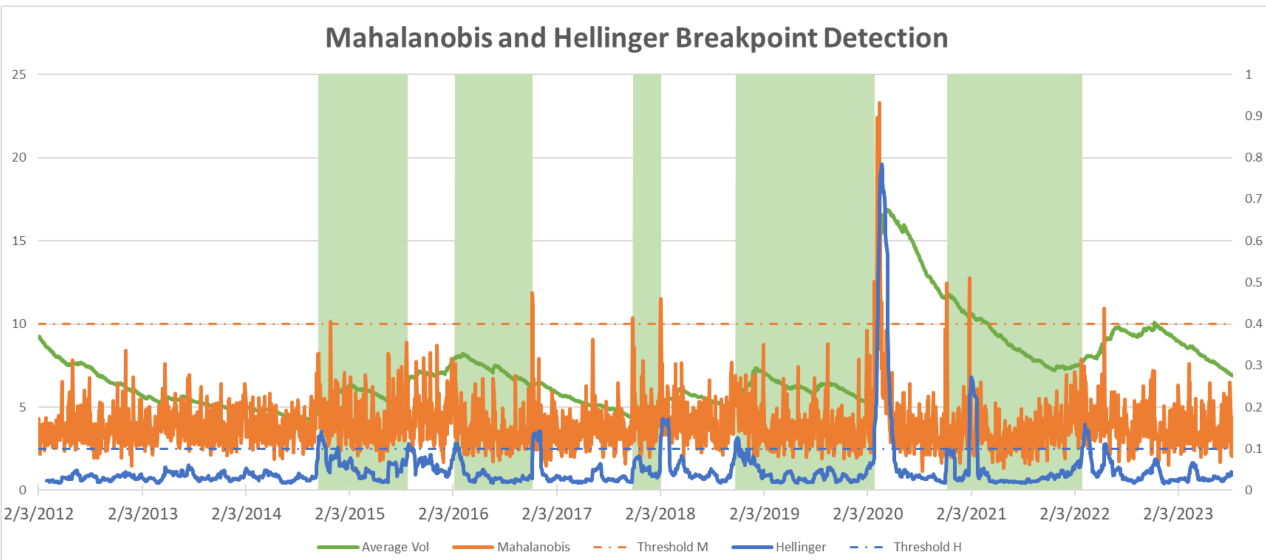

Figure 3: The 10 regime periods. On top of the Mahalanobis and Hellinger metrics, the average volatility of the markets shows, for each period, whether it corresponds to a high or a low volatility regime.

Conclusion

We used both the Mahalanobis metric and the Hellinger distance to identify breakpoints in the statistical behavior of the US stock market industry indices between 2012 and 2023. The Mahalanobis metric measures, each day, how the vector of industry sector returns deviates from the mean and covariances calculated on a historical series ending the previous day. The Hellinger distance measure, each day, how dissimilar are the current mean returns and covariances with respect to what they were a month before. Breakpoints are identified when one of these two metrics exceeds a threshold.

We found 10 breakpoints during this period, carving 11 periods of relatively stable statistics. We could identify several famous volatility surges, such as the COVID crisis in March 2020, but also at the beginning of February 2018 and other famous mini-crises. Interestingly enough, we also detected, thanks to the Hellinger criterion, periods where the volatility dropped, such as the end of 2016 or the end of 2020.

References

(1) Journal Article: Salhi, Khaled. Regime switching model for financial data: Empirical risk analysis. Physica A: Statistical Mechanics and its Applications, 461: 148-157, 2016.

(2) Journal Article: Mahalanobis, Prasanta Chandra. On the generalised distance in statistics. Proceedings of the National Institute of Sciences of India, 2 (1): 49–55, 1936.

(3) Internet: Wikipedia contributors. Mahalanobis distance. Wikipedia, The Free Encyclopedia, (accessed 31 October 2023): 2023.

(4) Journal Article: McLachlan, GJ. Mahalanobis Distance. Reson 4, 20-26, 1999.

(5) Journal Article: Hellinger, Ernst. Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen. Journal für die reine und angewandte Mathematik (in German), 136: 210–271, 1909.

(6) Journal Article: González-Castro, Víctor. Class distribution estimation based on the Hellinger distance. Information Sciences, 218: 146-164, 2013.

(7) Internet: Wikipedia contributors. Hellinger distance. Wikipedia, The Free Encyclopedia, (accessed 31 October 2023): 2023.

(8) Working Paper: Ang, Andrew, Timmermann, Allan. Regime Changes in Financial Markets, NBER working paper 17182, June 2011.