

Burkett, Matthew W.; Scherer, William T.; Todd, Andrew

Article

Are financial market states recurrent and persistent?

Cogent Economics & Finance

Provided in Cooperation with:

Taylor & Francis Group

Suggested Citation: Burkett, Matthew W.; Scherer, William T.; Todd, Andrew (2019) : Are financial market states recurrent and persistent?, Cogent Economics & Finance, ISSN 2332-2039, Taylor & Francis, Abingdon, Vol. 7, Iss. 1, pp. 1-21, <https://doi.org/10.1080/23322039.2019.1622171>

This Version is available at:

<https://hdl.handle.net/10419/245242>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Are financial market states recurrent and persistent?

Matthew W. Burkett, William T. Scherer & Andrew Todd |

To cite this article: Matthew W. Burkett, William T. Scherer & Andrew Todd | (2019) Are financial market states recurrent and persistent?, Cogent Economics & Finance, 7:1, 1622171, DOI: [10.1080/23322039.2019.1622171](https://doi.org/10.1080/23322039.2019.1622171)

To link to this article: <https://doi.org/10.1080/23322039.2019.1622171>



© 2019 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.



Published online: 11 Jun 2019.



Submit your article to this journal [↗](#)



Article views: 291



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)



Received: 07 February 2019
Accepted: 09 May 2019
First Published: 09 June 2019

*Corresponding author: Matthew W. Burkett, Systems and Information Engineering, University of Virginia, Charlottesville, USA
E-mail: mburky@gmail.com

Reviewing editor:
Vassilios Papavassiliou, University College Dublin, Dublin, Ireland

Additional information is available at the end of the article

FINANCIAL ECONOMICS | RESEARCH ARTICLE

Are financial market states recurrent and persistent?

Matthew W. Burkett^{1*}, William T. Scherer¹ and Andrew Todd¹

Abstract: Market participants often invoke the concept of discrete state when discussing financial markets. Bull market, bear market, depression, and recession are all terms that map to discrete market states. Mental models of how markets behave in each state and transition between states are then applied to decision-making. Implicit to that approach is the assumption that states are persistent and recurrent over time. This article seeks to formalize notions of discrete market states by proposing a parsimonious and innovative approach to segmenting periods of time into discrete states. The technique is demonstrated and evaluated in a series of case studies.

Subjects: Economic Theory & Philosophy; Mathematical Economics; Economic Forecasting; Finance; Investment & Securities

Keywords: portfolio design; regimes; discrete states; clusters; modern portfolio theory; Markov; states

1. Introduction

Financial markets are complex systems representing the global interaction of millions of individuals with widely varying objectives. The actions of each participant influence the state and behavior of the financial market system, which in turn influences the reactionary, future behavior of other market participants. The complex nature of such a system can result in abrupt changes in behavior, which may significantly impact participants (Farmer et al., 2012). Consider, as an example, the drastic change in return, volatility and correlation patterns exhibited in the global stock markets during the global financial crisis of 2008. Many participants suffered financial ruin during this period because they neither anticipated nor understood the scale to which the market behavior would change. Alternatively, others profited because they were able to anticipate the

ABOUT THE AUTHOR

The author, Matthew W. Burkett, is a PhD candidate at the University of Virginia in Systems and Information Engineering. He is a member of the Financial Decision Engineering (FDE) research group, which pursues areas of analytical research in the fields of finance and economics in pursuit of providing research that supports improved decision-making in the field as well as quantifying the underlying behaviors of markets and market mechanisms.

PUBLIC INTEREST STATEMENT

In this research article, a new technique is presented for classifying and describing periods of time in the financial markets. Market participants often think in terms of discrete states when discussing financial markets. Bull market, bear market, depression, and recession are all terms that represent discrete market states. Mental models of how markets behave in each discrete state and when transitioning between states are presented. This article seeks to formalize notions of how market states are defined by proposing a parsimonious approach to segmenting periods of market time into discrete states and offers methods for evaluating the resulting system of discrete states. The technique is demonstrated and evaluated in a series of case studies.

changes in market behavior and respond to the prevailing market conditions. The impact that market behavior can have on the well-being of investors has motivated observers to understand and better describe the nature of markets.

One approach that remains popular to classify, model and describe the behavior of financial markets is the idea of “state”. While in finance this concept goes by many names, such as regime, cycle, phase or market, the underlying concept is the same—a subset of the greater whole than is described in terms of key characteristics the values of which differentiate one state from others. There are several reasons that “states” are well suited for applications in the financial markets. First, the notation is natural and intuitive. Individuals already think of the financial market and economy in terms of states whether they are aware of it or not. One of the most well-known examples of state-based applications in the economy pertains to the business cycle. A robust economy is known as an *expansion*; however, a declining or contracting economy is a *recession*, a distinction made ex post facto. Whether the economy is classified as a recession or not is an example of defining a financial market state. Many individuals and researchers think of the equity markets as either advancing or declining, or to use the parlance of the financial markets, to be either a “bull market” or a “bear market.” To classify market periods in a binary fashion is to already consider financial markets in terms of discrete states.

The highly influential work on the subject, Hamilton (1989) introduces an original approach for modeling the business cycle by examining changes in regime using a discrete-state Markov process. Using a four-state model, Hamilton showed the change in underlying regime parameters as an economy shifted from one discrete state to another (Hamilton, 1989). When modeling and analyzing a system using a discrete state-based framework, there are several important questions that must be addressed and answered. One of the most important of these asks whether the discrete states of this system are recurrent and persistent.

This paper proposes an innovative approach for defining and evaluating states in the financial markets to investigate the question of whether market states are recurrent and persistent. The approach is presented according to the following framework:

- (1) Introduce an innovative and parsimonious feature set on which to make state classification decisions
- (2) Create market states using clustering algorithms
- (3) Calculate the probability of membership in each state for observation in the hold out set
- (4) Evaluate the quality of states by examining the state transition dynamics and feature similarities
- (5) Examine the results of the innovative approach through case studies

This approach is unique compared to state-based modeling research due to the dynamic, multistate nature of the approach. Additional, states are defined according to the underlying data without consideration to the characteristics a state should have in advance.

2. Data

The analysis incorporated within this paper is based on daily price and yield data for seven investment assets representing the major US equity indices and bonds from 1979 to 2018. Four domestic equity indices of varying market capitalization are included: S& P 500, NASDAQ Composite, Dow Jones Industrial Average, and the Russell 2000. The three bond products, which represent both public and corporate bonds, are: 10-year Treasury, Moody’s AAA, and Moody’s BAA.¹

3. Market state measures

The intended use of any state-based model must be a primary consideration when selecting the feature set from which states will be created and the associated feature set must be selected with this

objective in mind (Scherer, Adams, & Beling, 2018). It needs to capture the key characteristics that differentiate states from one another in a meaningful way. Since the intended use of this methodology is to offer insight into the intelligent design of investment portfolios, the choice of features must support the intended use by providing information about the metrics of interest. Investment portfolios are commonly evaluated in terms of how well the portfolio balances the tradeoff between potential risk and potential reward. Harry Markowitz (1952) proposed a formal, scientific approach to address this problem in the seminal work, *Portfolio Selection* (Markowitz, 1952). Markowitz introduces a mean-variance optimization (MVO) model that yields the optimal portfolio for every expected return level possible for a given universe of assets. Markowitz believes that the optimal portfolio for any level of expected return is the one with the least risk, which is commonly measured in term of volatility, specifically standard deviation. His MVO model has two parameters—a growth vector of the expected returns, r , of the assets in the universe of investments and the covariance matrix of asset returns, Σ . The growth vector affects the expected return aspect, and the covariance matrix affects the expected risk aspects represented in terms of volatility, or uncertainty, because the risk of a portfolio is a function of both the amount allocated to each asset and co-movement of the assets.

The parameters of Markowitz's MVO model serve as the foundation for our features selected to define market states; therefore, covariance acts as a measure of risk, and expected growth acts as a measure of return. Clustering methods use this feature set to segment the financial markets into states. These features—expected growth vector and covariance matrix—are multidimensional measures that each need to be reduced to a single number to be used by the clustering algorithms. The intended application of a state-based model determines the features by which states are defined. Researchers have used a variety of features in the state definition process: Neftci (1984) used unemployment rates (Neftci, 1984), Hamilton (1989) used GNP data (Hamilton, 1989), Ghezzi and Picardi (2003) applied dividend growth rates (Ghezzi & Piccardi, 2003), Badge (2012) examined macro-economic factors (Badge, 2012), Owyang and Ramey (2004) used elements of monetary policy (Owyang & Ramey, 2004), Guidolin and Timmerman (2007) uses equity and bond returns to classify market cycles as “crash”, “slow growth”, “bull” and “recovery” states (Guidolin & Timmermann, 2007), and Whitelaw (2000) uses expected returns and volatility (Whitelaw, 2000). This research presents a unique perspective in terms of the feature set used to define states by considering elements of the covariance matrix. The covariance matrix reflects the co-movement of assets and, as Markowitz's MVO model illustrates, impacts the volatility characteristics of a portfolio. The facts that this research presents support intelligent portfolio design and the influence of the covariance matrix on the overall portfolio volatility were the reasons for using a covariance-based feature in the state definition process.

The first feature measures the expected return, or growth, vector over z periods of time and condenses the aggregate growth into a single number by considering the expected returns of all the individual component assets and representing the expected growth vector as a magnitude of growth for all the assets. The return of z periods is calculated for each asset and summed for all n assets in the set, r . The square root of the sum of squares is taken to condense the performance to a single number. The look-back parameter z can be adjusted at the discretion of the researcher based on the level of data granularity and the intended use of the analysis results. Equation 1 shows that the magnitude of growth is calculated using the Euclidean distance from the zero-return origin and represents the intensity of asset returns over the period of observation.

$$m(r^{(z)}) = \sqrt{\sum_{i=1}^n r_i^2} \quad (1)$$

As shown in Equation 1, the magnitude measure, m , does not reflect if the assets are moving in an increasing, positive manner or in a decreasing, negative one so a directional component, d , is introduced to account for the market direction. The addition of the component d is intended to capture the overall directional trend of asset returns in r . To capture the prevailing directional trend of prices in r , the mean of returns of r is determined. Equation 2 shows that the variable d is directional

flag indicating whether the mean return is positive or negative by assuming a value of 1 or -1, respectively.

$$d(\mathbf{r}^{(z)}) = \begin{cases} 1 & \text{if } \frac{1}{n} \sum_{i=1}^n r_i^{(z)} > 0 \\ -1 & \text{otherwise} \end{cases} \quad (2)$$

The product of results from Equations 1 and 2 lead to the return growth measure, $g(\mathbf{r})$, shown in Equation 3. The growth measure is a discontinuous function that experiences “jumps” in the plot of the function when the sign of d changes.

$$g(\mathbf{r}^{(z)}) = d(\mathbf{r}^{(z)}) * m(\mathbf{r}^{(z)}) \quad (3)$$

Each of the subcomponents of the expected growth measure are shown in Figure 1(a). The magnitude value is plotted, and the value of the directional flag is overlaid in red. The addition of the directional flag results in gap between the elements of the subpopulations which have different values as shown in Figure 1(b).

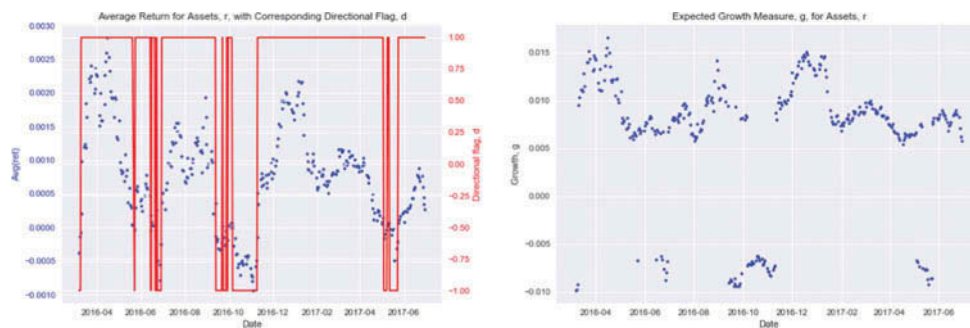
The second feature, which measures the covariance matrix of assets in the investment universe, reduces the n by n covariance matrix to a single number that can differentiate between periods of time in terms of the co-movement of assets. A common approach for quantifying a matrix, particularly a covariance matrix, is by calculating the distance between a base covariance matrix, \mathbf{A} , acting as a reference point and a covariance matrix of interest, \mathbf{B} . The base matrix may be selected for a variety of reasons, such as representing values of interest or an initial, starting state or values simply selected at random. This research applies a technique presented by Förstner (1999), which also calculates the distance to a base matrix, and is presented in Equation 4 (Forstner & Moonen, 1999).

$$d_F(\mathbf{A}, \mathbf{B}) = \sqrt{\sum_{i=1}^n \ln^2 \lambda_i(\mathbf{A}, \mathbf{B})} \quad (4)$$

where $\lambda_i(\mathbf{A}, \mathbf{B})$ is the joint eigenvalue of the matrices.

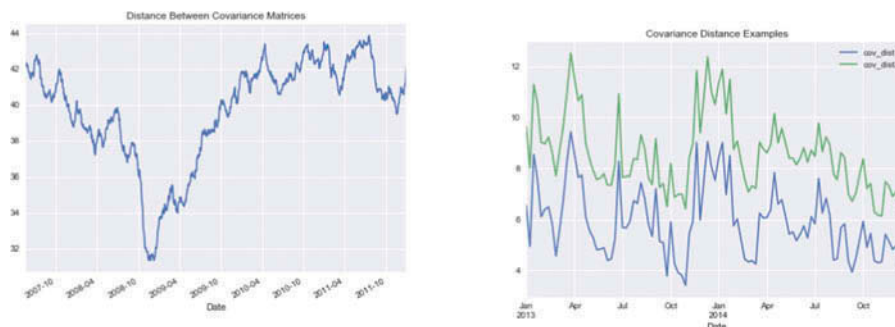
As mentioned previously the calculation for d_F is made relative from a static, base covariance matrix, \mathbf{A} , to a matrix of interest, \mathbf{B} . Since all distances are affected by the choice of baseline matrix \mathbf{A} , there is a concern regarding the degree to which this selection affects the distance calculation. This question was examined by taking two periods² that represent distinctive periods of market behavior and plotting the distance between the same \mathbf{A} matrix. Two baseline matrices, \mathbf{B}_1 and \mathbf{B}_2 , are calculated using all observations from Q1-Q2 2008 and Q3-Q4 2008, respectively. The \mathbf{A} matrix incorporates a 22-day lookback period. This means that for each point in time t from the examination period (2013–2014) the previous 22 observations are used to calculate the covariance matrix. The distance between \mathbf{A}_t and \mathbf{B}_1 and then between \mathbf{A}_t and \mathbf{B}_2 is calculated. The examination window shifts one period and \mathbf{A} is

Figure 1. Expected return feature, $g(\mathbf{r})$.



(a) Average return for assets in r . The directional flag, d , indicates whether the average return is greater than zero or not. (b) The feature $g(\mathbf{r})$ combines two component features: a directional component $d(\mathbf{r})$ and a magnitude vector $m(\mathbf{r})$.

Figure 2. Covariance feature.



(a) An example of the covariance distance described in Equation 4. (b) Impact of different baseline matrices (2008: Q1-Q2, Q3-Q4) in covariance distance calculation.

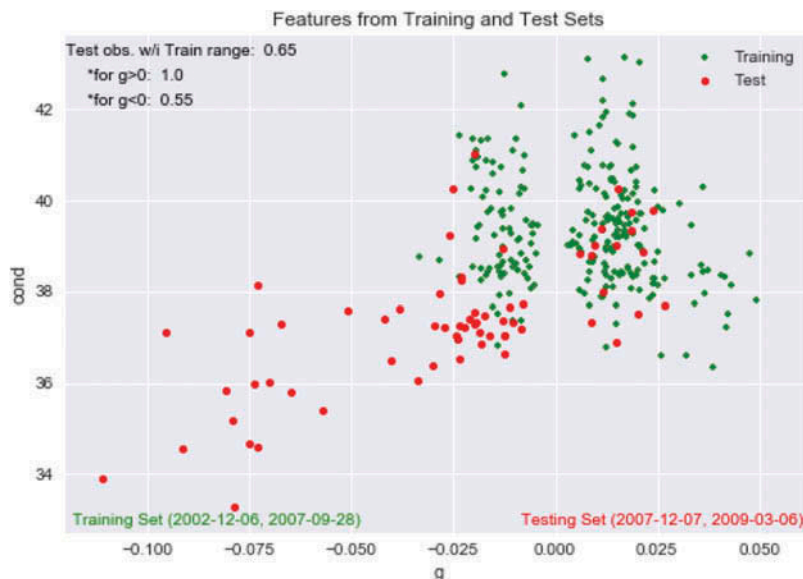
recalculated using the observation from time $t + 1$ and the previous 22 periods. The distances between A_t and B_2 are then recalculated and plotted. As shown in Figure 2(b), the selection of a base covariance matrix does affect the distance calculation; however, this effect is constant between points and can also be viewed as a vertical shift of distance plots. The constant difference in distance reduces the uncertainty and variability associated with this calculation as a quantifiable measure of the covariance matrix.

4. Techniques for determining market states

4.1. Assignment of the training set and cluster creation

The overall dataset must be subdivided into two smaller subsets: a training set which will be used to create the market states and a holdout, testing set which will be used to evaluate the state schema created on the training data. The selection of training and testing periods is a key step in the modeling process. Bailey et al. analyzed the problem of overfitting back-tested results in the context of investment returns. The paper presents that traditional statistical techniques designed to prevent or at least minimize regression overfitting are ill-suited and unreliable when applied to investment models and suggests the need to apply non-standard methods (Bailey et al., 2016). For the clustering methodology to be robust enough to be applied in periods beyond the training period and thereby prevent overfitting, the range of values across the list of features in the training set should encompass the range of values in the testing set. Figure 3 illustrates the observations

Figure 3. A comparison of observations from the training and test sets. Periods from the test region should fall within the bounds of training set observations. If an insufficient number of observations fall within these bounds then the training and testing regions should be reconsidered.

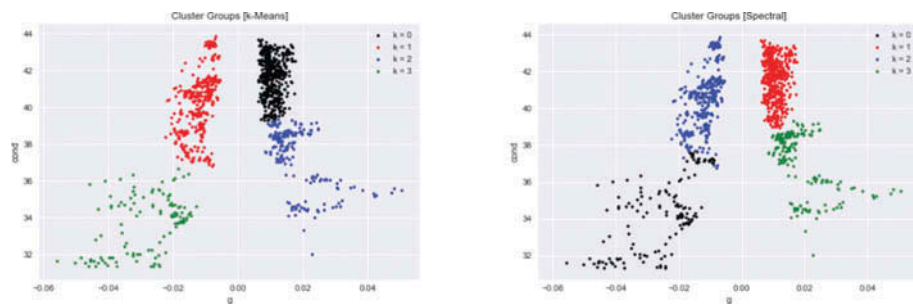


from the training and testing sets to show whether the training region is representative of the testing region, which speaks to the robustness of the modeling approach. For this example, which includes observations of domestic market equities, 43% of all observations from the testing region (12/7/2007–3/6/2009) fall within the range of values from the training set (12/6/2006–9/25/2007). Differences are also examined based on the gaps created by the d component of the growth vector, g , calculated in Equation 2, which determines the sign of the vector. The percent of testing values that fall within the range of training values for positive and negative g values is 45% and 29%, respectively. The boundaries of the training and testing regions shown in this example should be reconsidered because observations with low covariance distance values and large growth magnitudes in the testing set fall outside the boundaries of observation within the testing set. The modeling schema presented does not consider data points within this range of values and is therefore not correctly calibrated to address such data points properly.

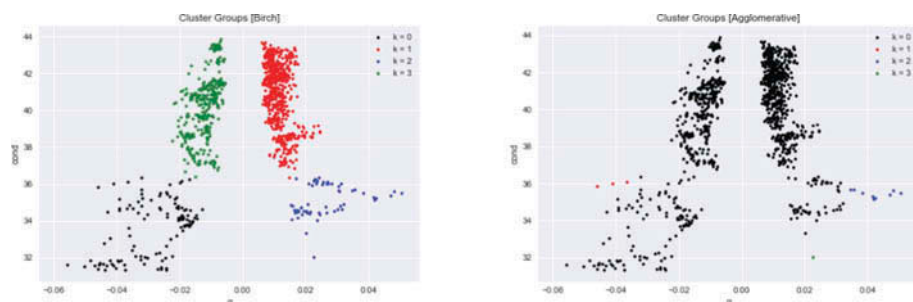
Each observation can be quantified in terms of the two features of interest described in the previous section. Using these features, states within the financial markets can be created by grouping similar observations. Many techniques exist for creating groups, but this research applies unsupervised clustering algorithms from the field of machine learning. Differences in the groups created by the various algorithms is a function of the methodology applied by the algorithm as well as the inputs and even randomness (Romesburg, 2004). Many unsupervised learning approaches could be used to create states, such as DBSCAN, spectral methods, or Birch's; however, this research presents results created using the k-means algorithm. The k-means algorithm was selected because it is a well-known and commonly used algorithm that produced similar clustering results in other potential methodologies as shown in Figure 4.

Prior to the application of a clustering algorithm, the issue of scale must be addressed by placing all points on the same scale. Features exist on a variety of size scales, and cluster assignments are biased towards the feature with the larger scale. To account for the differences in scale of the

Figure 4. Comparison of clustering methods results for $k = 4$ configuration.



(a) Results for k-Means. Similar to those in 4b and (b) Spectral clustering results. Similar to those seen in 4a and 4c.



(c) Results for Birch's Method. Similar to results (d) Agglomerative Method results are unlike others. Algorithm created one dominant cluster.

growth and covariance features, the data used to create the clusters must first be normalized. A common method for normalizing data sets is to calculate the number of standard deviations an observation is from the mean and represent the value in those terms according to Equation 5.

$$x^* = \frac{x - \bar{x}}{s_x} \tag{5}$$

where \bar{x} and s_x are the sample mean and standard deviation, respectively, of the population of the feature x .

Once the data has been normalized, a clustering algorithm can be applied without bias to feature size. Examples of results from four (4) algorithms are presented in Figure 4 to show a sample range of results based on technique. All four methods required k as an input, so a value of $k = 4$ was provided to each. The optimal numbers of clusters k will vary based on the problem. For the purposes of this article, a value of $k = 4$ was selected to provide a clustering solution that is general enough demonstrate the overall concept while providing enough detail as to illustrate pertinent differences in the clustering solutions presented. The methodology for selecting the value of k falls outside the scope of this article; however, this question lies at the center of creating an effective clustering methodology. Similar results were achieved by three of the four algorithms; however, one method—Agglomerative Method—produced a single, dominant cluster comprising the vast majority of observations with the remaining clusters being sparsely populated (Figure 4(d)). The results for the Agglomerative method show that it is not well suited in this case and was removed from consideration; however, given the similarities in results of the other three methods, one can be used as an illustrative example. The k-means algorithm is a well-known, commonly used method that produced results like those of the other two methods and therefore will be used as the example algorithm.

Applying a clustering algorithm to the training set, each daily observation consisting of the n -day growth vector magnitude and n -day covariance matrix distance pair is assigned to a specific state. Figure 5 shows the results of this sample clustering schema, which indicates the state to which each observation from the training set was assigned. Daily observations from the training set are assigned to one specific state because these cluster assignments define the framework for the overall state-based schema and subsequently acts as the foundation on which all future clustering decisions are based. Observations from the training set are not assigned absolutely to a cluster.

Figure 5. The assignment of observations to states in the training using the k-means algorithm.



Assignments from observation in this set are made probabilistically by considering the likelihood of membership in each of the various clusters.

4.2. Probabilistic assignment of observations in the testing set

Unlike observations in the training set, which are assigned to a specific cluster (see Figure 5), the observations in the testing set are not explicitly assigned membership in a particular cluster. Instead observations from the testing set are given a likelihood of membership a state based on similarities to the “typical” state member, which is expressed as the state’s centroid.³ Therefore the distance between an observation x^* and the centroid for cluster k , $x^{*(k)}$, can be calculated by the formula:

$$d_k(x^*, x^{*(k)}) = \sqrt{\sum_{j=0}^{n_j} (x_j^* - x_j^{*(k)})^2} \quad (6)$$

where j is the feature, and n_j is the total number of features used to describe the state.

Throughout this paper, two features—growth magnitude vector (g) and covariance distance (cov)—are used as the features of interest to create states. The specific distance formula for this example is:

$$d_k(x^*, x^{*(k)}) = \sqrt{(x_g^* - x_g^{*(k)})^2 + (x_{cov}^* - x_{cov}^{*(k)})^2} \quad (7)$$

The distance measure d_k from Equation 7 serves as the sole input parameter for determining state probability membership; however, the probability that an observation is a member of any given state is inversely related to its distance from the cluster’s centroid. The closer an observation is to a cluster centroid the greater the degree of similarity to the typical member of the cluster. In terms of the variable d_k , an observation with a lower distance value to cluster k has more in common with the members of that cluster than an observation with a higher d_k and therefore has a higher probability of membership in that state. The inverse d_k ratio must be normalized as shown in Equation 8 to transform the ratio into a membership probability. The relationship between distance to a cluster can be expressed as:

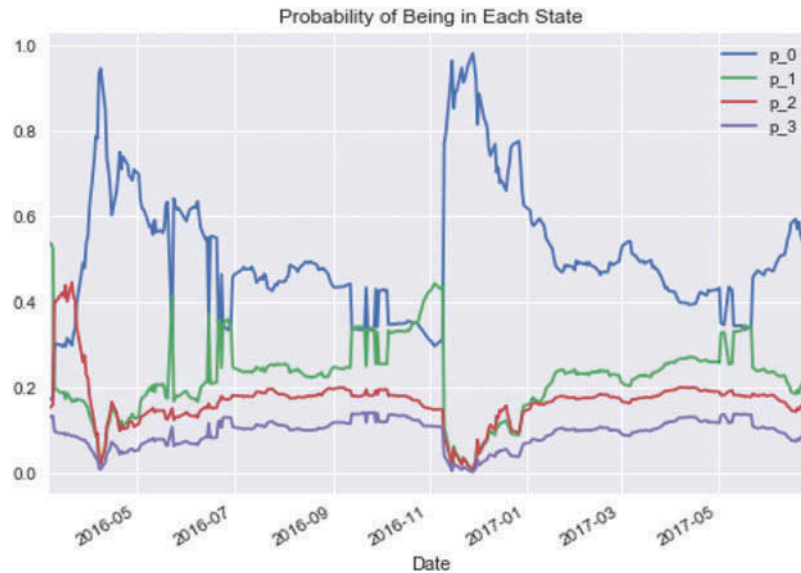
$$p_k = \frac{d_k^{-1}}{\sum_{i=0}^n d_i^{-1}} \quad (8)$$

Using the relationship from Equation 8, the probability of membership in each state for observations in the testing set is shown in Figure 6.

Using this methodology, each observation is given a probability of membership to every state. While such an approach has its benefits, chief among them is consideration to low probability events; however, this can also introduce noise into the analysis in the form of low probability membership events. One approach to address this issue is through the introduction of a level of confidence for state membership. This can be accomplished through the introduction of confidence ellipsoids, which is the multivariate (in this example, a bivariate case is presented) analog to a statistical confidence interval. Like a confidence interval, confidence ellipsoids relate to the underlying population parameters. For example, a 95% confidence ellipsoid would not necessarily be expected to encircle 95% of the state observations; however, if repeated sampling from the underlying distribution were performed, the population mean would be expected to fall within the ellipsoid 95% of the time. For the purposes of this research, the ellipsoid acts as a quantifiable threshold for deciding whether an observation is eligibility for considerable as a member for a state (Draper & Smith, 2014; Friendly, Monette, & Fox et al., 2013).

The confidence ellipsoid is a function of four variables: 1) the magnitude of the major axis, 2) the magnitude of the minor axis, 3) the scale of the ellipsoid (the level of confidence the ellipsoid

Figure 6. Using the clusters created using training data, a probabilistic estimate of cluster membership is made for the observations from the holdout, test set.



represents), and 4) angle of orientation about the axis. The first two variables—magnitude of the major and minor axis, or height and width—are a function of the variance. High variance indicates that the data is widely distributed, and the mean is not well established, so the ellipsoid will be extended further in the direction of the high variance distribution. The basic equation in Equation 9 for an ellipsoid is applied to generate a shape with the appropriate proportions given the variance of each feature, a and b . The a and b axis represent features, which in the case of this example would be the growth vector and covariance distance, respectively. Large ellipsoids represent states with a wide range of values in one or both features, which would indicate that the state is more general than a state with a smaller ellipsoid and more concentrated feature set of its members.

$$f(a, b) = \left(\frac{a}{\sigma_a}\right)^2 + \left(\frac{b}{\sigma_b}\right)^2 = s \tag{9}$$

The third variable—confidence scaling factor of the ellipsoid—increases the size and by extension the level of confidence the results represent; however, the estimations about confidence need to be made with respect to a specific, underlying statistical distribution. A Chi-Squared distribution was selected because the terms $\left(\frac{a}{\sigma_a}\right)^2 + \left(\frac{b}{\sigma_b}\right)^2$ in Equation 9 represent an independent sum of squares distribution, and the sum of squared Gaussian observations is distributed according to a Chi-Squared distribution (Johnson & Kotz, 1970). Using this distribution assumes of normality within the feature set, which may not hold true for all instances; however, the goal of this application is to provide a heuristic for assigning observations and therefore an approximation of statistical confidence is sufficient for this purpose. This scaling factor can be introduced by replacing s in Equation 9 with the appropriate value from the Chi-Squared distribution that represents the appropriate level of confidence. Three commonly used values ($df = 2$) are: 90% 4.605, 95% 5.991, and 99% 9.210. An increase in the confidence level corresponds to a larger S value, which in turn corresponds to a larger ellipsoid. All points which satisfy Equation 10 meet the criteria for membership in the state at the given level of confidence.

$$\left(\frac{a}{\sigma_a}\right)^2 + \left(\frac{b}{\sigma_b}\right)^2 \leq S \tag{10}$$

The first three variables defined the shape and size of the ellipsoid. The fourth variable specifies the orientation of the ellipsoid. Just as the variance of each of feature affects the magnitude of the

major and minor axis of the ellipse, the direction of the variance as defined by the covariance matrix impacts the orientation of the ellipsoid as well. The eigenvalues of the covariance matrix— $v(a)$ and $v(b)$ —represent the variance of data in the direction of the eigenvectors. Using Equation 11, the angle of orientation, θ , between the major axis and the x-axis can be calculated.

$$\theta = \arctan\left(\frac{v(a)}{v(b)}\right) \tag{11}$$

Deriving the orientation angle to the x-axis from Equation 11 and applying it to the ellipsoid created through Equation 10 can be used to create confidence ellipsoids for each example result shown in Figure 4(a). The resulting cluster assignments with corresponding confidence ellipsoid overlays are illustrated in Figure 7.

The ellipsoid boundaries calculated in Equation 10 and shown in Figure 7 can be applied to the test set to determine the threshold for which states an observation can be probabilistically assigned and which states are removed from consideration. Figure 8 shows the observations from the test set with the confidence ellipsoids based on the states from the training set. Observations that only fall within a single ellipsoid would be assigned to a single state, observations that reside in overlapping ellipsoids would be probabilistically assigned to states as described in Equation 8. In the next section, we discuss how to evaluate the states created using this approach.

5. Methods of evaluating market states

Having outlined a methodology for measuring key market characteristics, defining market states based on those key characteristics, and probabilistically assigning future periods to each of these market states, the results of this methodology must be evaluated to determine whether they offer value through the application of this state-based approach. The evaluation criteria revolve around the ideas of recurrence and persistent behavior. Both are essential to the practical application of any results derived from this method. First, for any insight derived from this approach to be useful, the state must recur. Otherwise, there would be no opportunity to use any insight gained. Secondly, knowing that a market state will recur at some time in the future is a necessary but not sufficient condition for adding value. There must also be a degree of insight into the behavior of the market while in that state.

Visual inspection of the transition behavior in Figure 5 suggests that observations are capable of 1) leaving a state indicating that it is *non-absorbing* in nature and 2) return to any state

Figure 7. Cluster results from a $k = 4$ schema with an overlay of two-standard deviation confidence ellipsoids.

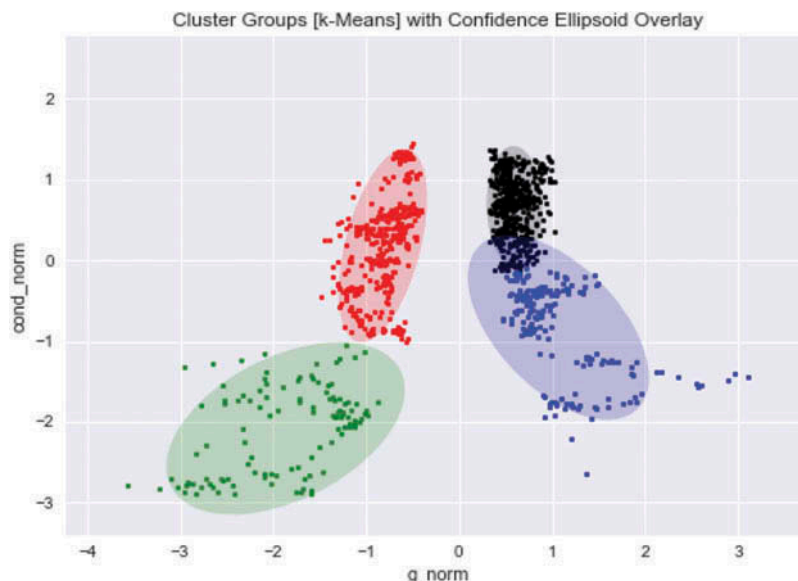
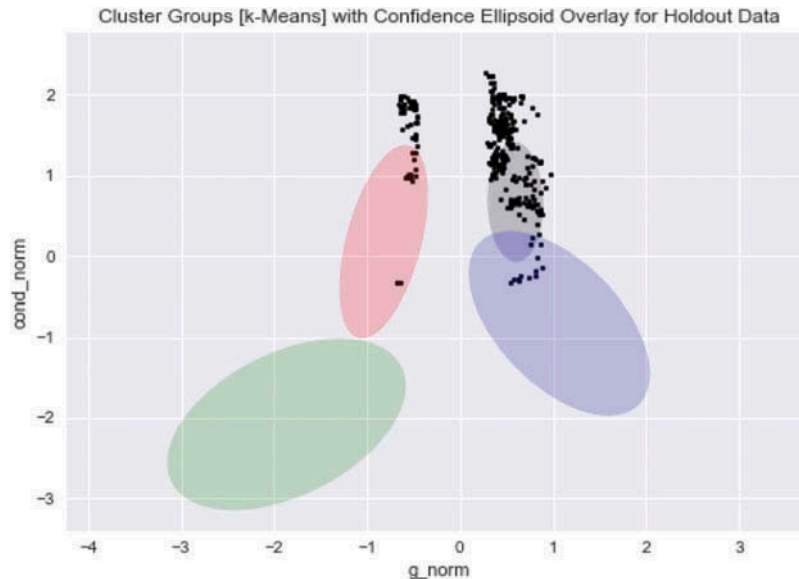


Figure 8. Observations from the holdout, test sets are plotted against the two-standard deviation confidence ellipsoids for clusters to estimate the cluster membership for the holdout observations.



after transitioning from it indicating *recurrent* behavior. Both are necessary characteristics for this research because to examine the transition dynamics in a scientific and objective manner, the system of market states should be modeled as a Markov process. Such a model determines the probability of transitioning to future states j is governed only by the current state i , and any state visited prior, denoted as k , does not affect the transition dynamics of the system (Bertsekas, 2002).

A transition probability matrix is a method of representing the transition behavior of a system. The matrix P in Figure 9(a) represents the probability a system being in state j conditional on visiting i in the previous time stage, p_{ij} , for all combinations of states $(0, n)$, where $p_{ij} = P(x_{t+1} = j | x_t = i)$.

Visual inspection of Figure 5 demonstrates that each state is visited more than once suggesting that if a system leaves any state, it can return to that state. This finding of recurrent behavior can be reinforced by examining the system’s transition probability matrix from which the long-run, steady-state probability matrix can be calculated.⁴ Using the example transition matrix shown in Figure 9(b), steady-state probabilities for the system are [0.098 0.544 0.053 0.304]. All probabilities are greater than zero showing that all states can be reached at some point in future. If a state was not recurrent, the steady state for it would be zero (Bertsekas, 2002). This characteristic is important in modeling the financial markets from a state-based perspective because the value added from such a model is predicated by the expectation that the system return to a given state, which in turn provides insight into the behavior of the system. The transition probability matrix and data used to create it can be used to examine the existence of the Markov property.

Figure 9. Transition probability matrices for data in Figure 7.

$$P = \begin{pmatrix} p_{0,0} & p_{0,1} & \dots & p_{0,n} \\ p_{1,0} & p_{1,1} & \dots & p_{1,n} \\ \vdots & & & \\ p_{n,0} & p_{n,1} & \dots & p_{n,n} \end{pmatrix}$$

(a) Generalized transition probability matrix

$$P = \begin{pmatrix} 0.91 & 0.00 & 0.09 & 0.00 \\ 0.00 & 0.95 & 0.00 & 0.05 \\ 0.15 & 0.02 & 0.84 & 0.00 \\ 0.01 & 0.09 & 0.00 & 0.90 \end{pmatrix}$$

(b) Matrix for training set example

The existence of the Markov property can be examined using a frequency-based approach to determine if being in state k prior i had an effect on the transition probabilities to state j . The χ^2 test was used to determine if there is a statistically significant difference in transition probabilities based on the prior state k . All i, j combinations should be examined to determine the degree to which the system as a whole exhibits the Markov property (Devore, 2015; Scherer & Glagola, 1994).

Table 1 shows how to structure the χ^2 test for examining the Markov property. This test uses the matrix P , only the p_{ij} values where $i = 1$ and $j = 1$ are included. Separate tests must be performed on each i, j pair. The sample is divided into subsets depending on the state k visited prior to being in i and is represented by a row in the table. For a subset to be eligible for the χ^2 test, the expected number of observations must be equal to at least five (5) records. Otherwise, the row should be excluded from consideration (Devore, 2015). In this example, two rows ($k = 1, 3$) meet the minimum expected number of observation requirements and may be included in the calculation.

The χ^2 test is structured to examine the null hypothesis that equality exists among the transition probabilities given prior k state values, i.e. there is evidence of the Markov property. The statistical test found a p-value of 0.13 for the transitions for $p_{1,1}$; therefore, the test fails to reject the hypothesis that the Markov property exists for any level of significance less than 0.87. The testing should be repeated for all p_{ij} values in the same manner.

A second method to evaluate the efficacy of states created is to examine the persistence of the state characteristics between the training set on which the states were created and the testing set on which the states are evaluated. Specifically, the two features used to create the states—growth vector and covariance distance—are examined on a state-by-state basis using membership from the training set and distributions of feature characteristics are created. These distributions are then compared to those from the training set to determine if the features are statistically similar. Since observations from the testing set are probabilistically assigned to each state, testing set observations are assumed to be members of the state to which it has the highest probability of membership. The two-sample Kolmogorov–Smirnov (K-S) test compares two datasets to determine if they are statistically different. The K-S test was selected because it is non-parametric and makes no normality assumptions regarding the distribution of the data sets. The distributions of feature values for each feature-state pair are compared using the K-S test statistic, which is based on the point of maximum deviation between the cumulative distribution function of the training and test sets. Take as an example the growth vector for State 1. The distribution of values for both sets is shown in Figure 10(a), and the associated cumulative distribution from which the K-S test statistics is derived is shown in Figure 10(b).

$$D_{n,m} = \sup_z |F_1(z) - F_2(z)| \tag{12}$$

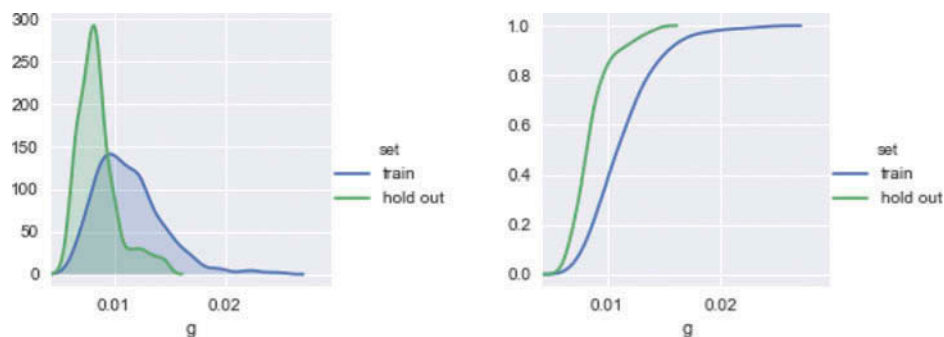
where $F_1(z)$ and $F_2(z)$ are the cumulative distribution being compared consisting of n and m observations, respectively, and \sup is the supremum function.

Using the cumulative distribution, the test statistics, D , is calculated to be 0.508 using Equation 12, which equates to a p-value of 0.0001. The null hypothesis that the two samples come from the same underlying distribution is rejected (Hazewinkel, 2001). This statistical result is supported by a visual examination of the distributions in Figure 10(a). The distributions of feature values, and by

Table 1. Result summary for χ^2 test of Markov Property: $p_{1,1}$

k	$p_{1,1} k$	$0_{j=1}$	$0_{j \neq 1}$	$n_{1,k}$	$E[\text{Obs}]$	Eligible?
0	0.00	0	0	0	0.00	
1	0.96	574	24	598	564.83	X
2	1.00	1	0	1	0.94	
3	0.70	23	10	33	31.22	X

Figure 10. Kolmogorov–Smirnov (K-S) test: growth vector.



(a) The distribution of growth vector values for training and hold out testing set for one state. (b) Cumulative distribution of growth values, which is used to create the K-S test statistic.

extension the behavior of the population in terms of this particular feature of interest, are statistically different between the training and test sets.

A third method for evaluating the results of clustering examines whether the transitional dynamics of the schema are robust. More specifically, it tests whether the transition behavior of the states from the training set as quantified by a transition probability matrix persists to states in the testing, hold-out set. As previously discussed, observations from the training set are members of states to a certain degree of probability as contrasted to states in the training set from which observations are assigned to states absolutely. To address this issue, observations are assigned to the nearest, highest-probability state from which the testing transition probability matrix is calculated. The comparison between the two matrices is made using the χ^2 test ((Bartlett, 1951; Billingsley, 1961)). This approach uses the state transition probabilities, p_{ij} , from the training set combined with the number of actual observations in the state i from the testing set to determine the expected number of observations in state j . A statistical test is conducted on the differences between the expected number of observations and the actual number of observations to determine whether the differences are statistically significant.

5.1. Results

The results of this methodology are presented in the form of three case studies which employ empirical data to demonstrate the outcome if this approach been applied to various historical periods. The bounds for each case are detailed in Table 2. A basket of domestic stock indexes and bonds⁵ was selected as the financial assets of interest based on overall notoriety and significance to the financial markets.

5.2. Case study 1: roaring 1980s & 1990s

The 1980s and 1990s represented a bullish period in domestic financial market history. The US experienced a strong economic recovery and sustained period of expansion that led to a steadily advancing domestic stock market free from periods of sustained negative performance, and bonds

Table 2. Time parameters for case studies

		Case 1	Case 2	Case 3
Training	Start	1/1/1980	10/1/2002	12/31/1979
	End	12/31/1994	10/1/2007	12/31/2015
Testing	Start	12/31/1994	10/1/2007	12/31/2015
	End	12/31/1999	3/10/2009	3/23/2018
Baseline Cov.	Start	1/5/1979	7/1/2002	1/5/1979
Matrix	End	12/31/1979	10/1/2002	12/31/1979

were paying high interest rates to compete with a strong equities market. The sustained period of economic and market growth was interrupted by a brief but notable market event occurred in October 1987 known as “Black Monday”, which represented a 22% single-day decline in the Dow Jones Industrial Average. The first case study examines this period using the first 15 years (1980 to 1994) as the training set on which to build and define the clustering scheme, and then the period from 1995 to 1999 to test and evaluate the results of the clustering. Each observation represents the 22-day growth magnitude and the covariance distance from the baseline covariance matrix as described in Figure 2. The reference covariance matrix is calculated using a data from a 22-day subset.

The features from the training and testing set are shown as a pairwise plot in Figure 11(a). The range of feature values for observations from the testing set are contained, more than 0.99 of the population, from known and observed data values contained in the training sample. The selection of the training and testing regions represents a crucial step in the modeling process. If values from the testing sample fall outside the range of feature values used to develop the cluster schema, the efficacy of the model comes into question as the data falls outside the scope of the modeling process as in the case shown in Figure 8. The training set presented in Figure 11(b) appears representative of the values in the testing evaluation set. The k-means algorithm is applied to the training feature set, and the cluster assignments are for the four-state configuration are displayed.

The results from this clustering schema are summarized using normalized values in Table 3. The clustering results distributed observations among states such that no single state contains most of observations nor is any state so sparsely populated as to make it lack representation of a substantive portion of the population. The *g* and *cov* values represent the centroids for each cluster.

The cluster-specific distribution of values for each feature is calculated using observations from the training set, which is compared with the distribution from the corresponding cluster-feature combination. The growth features for Cluster 3 do not persist between the training set and the testing set samples. The K-S test yielded a p-value of 0.007, for example, which conclusively rejects the hypothesis that the features were drawn from the same underlying distribution for the growth magnitude shown in Figure 12(a) with a high degree of certainty. Visual inspection of the figure confirms the differences between the distributions. Conversely, the distributions of covariance distance in Cluster 2 are similar and cannot be said to be statistically different. The K-S Test

Figure 11. Feature values and cluster assignment.

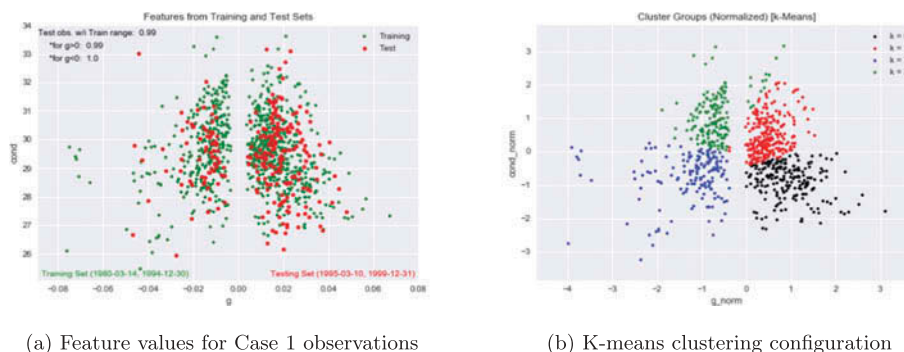
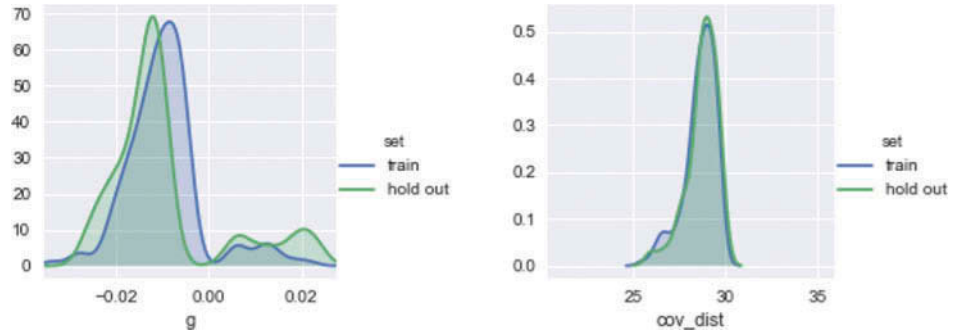


Table 3. Summary of Case 1 clusters (k = 4) from training set			
Cluster	No. of Obs.	<i>g</i>	<i>cov</i>
0	201	0.93	-0.89
1	255	0.48	0.50
2	157	-1.28	-0.75
3	159	-0.69	1.05

Figure 12. Kolmogorov–Smirnov (K-S) test: feature comparison.



(a) The growth vector values for training and hold-out testing set for Cluster 3. Statistically different. (b) Covariance distance values distributions for Cluster 2. Not statistically different.

produced a p-value of 0.669, which fails to reject the hypothesis that the two distributions are statistically similar as shown in Figure 12(b). Only two of the eight features were found to show statistical differences between the training and testing set values—growth features in Cluster 1 and 2. If the proportion of features that fail to pass the statistically similarity test falls below a threshold deemed significant as set by the user then the expectations of accuracy and predictability based on the clustering results should be tempered since the underlying feature behavior of populations from test and training sets is not statistically significant. The transition probability matrices, shown in Figure 13, were compared to one another to determine if the transition dynamics between states were persistent between the training and testing sets.

Visual inspections would suggest some similarities; however, the comparison must be made statistically using the χ^2 test. The results of the χ^2 test are shown visually in Figure 14. When χ^2 test value falls to the left of significance level threshold, such as 90% or 95% confidence level, then the test fails to reject the null hypothesis that the populations are similar. Any result that falls to the right of the significance level of interest has enough statistical evidence to reject the hypothesis that the populations are the same. The test produced a p-value of 0.13, which is close to the rejection region at 90% confidence level as shown in Figure 14 but does not fail statistically. Situations, such as this, are left to the discretion of the user when to reject the results or to accept them. The decision of selecting the appropriate level of significance for this test is left to the user.

5.3. Case study 2: financial crisis of 2008

The Financial Crisis of 2008 is considered by many to be the most severe financial market event since the Great Depression of the 1930s. It began with a catastrophic situation in the subprime mortgage market and expanded into a banking crisis resulting in the collapse of major institutions, such as Lehman Brothers. A global economic downturn ensued resulting in financial market prices falling worldwide. Because of the catastrophic impact this crisis, regulators around the world instituted policies to safeguard consumers and financial markets, such as Dodd–Frank Wall Street Reform and Consumer Protection Act and Basil III. The second case study examines this period to see how the state-based approach would perform during this watershed moment in financial market history. The training set used is the five years (10/2002 to 10/2007) prior to the

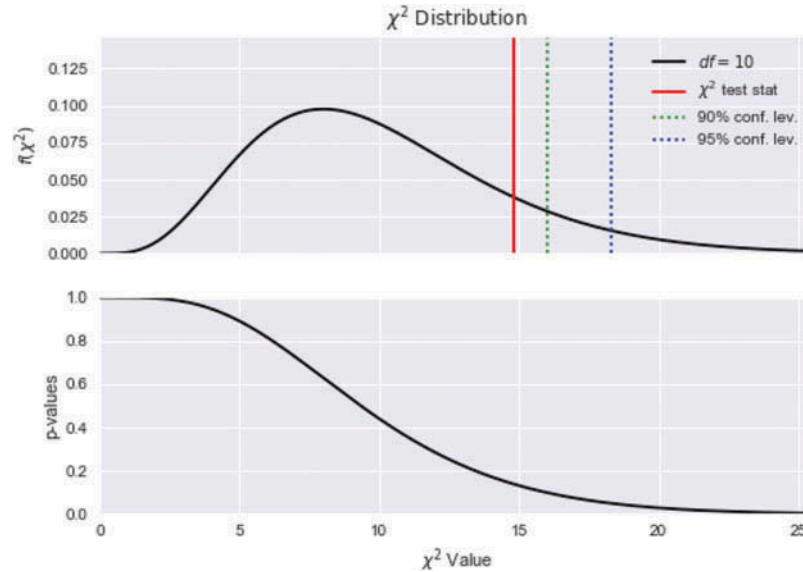
Figure 13. Comparison of transition probability matrices.

$$P_{train} = \begin{pmatrix} 0.76 & 0.18 & 0.04 & 0.02 \\ 0.12 & 0.71 & 0.02 & 0.15 \\ 0.08 & 0.06 & 0.73 & 0.10 \\ 0.01 & 0.17 & 0.15 & 0.65 \end{pmatrix} \qquad P_{test} = \begin{pmatrix} 0.71 & 0.23 & 0.03 & 0.01 \\ 0.13 & 0.75 & 0.04 & 0.06 \\ 0.09 & 0.11 & 0.62 & 0.16 \\ 0.00 & 0.17 & 0.26 & 0.55 \end{pmatrix}$$

(a) Training set probability matrix

(b) Testing set probability matrix

Figure 14. Results of χ^2 comparison of transition probability matrices. Fail to reject hypothesis that similar.

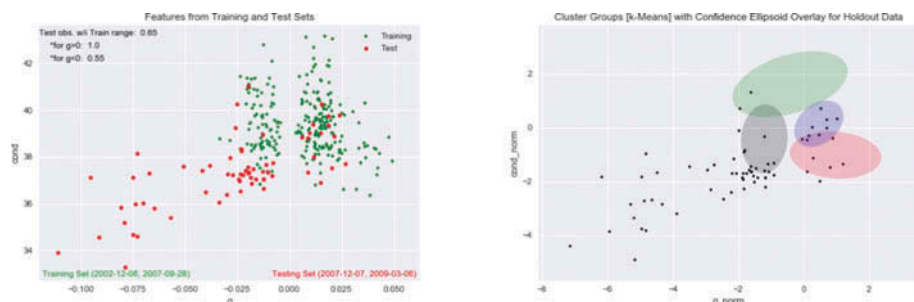


beginning of the financial crisis. The cluster schema created is tested on the market decline period from 10/2007 to 3/2009.

The importance of selecting a training set that is representative of the set on which the model impacts how robust a clustering model will be. The training period did not experience the magnitude of market events, particularly negative market events, that were experienced in crisis test set. Figure 15(a) shows that only 0.65 of the testing set falls within the bounds of the training set. Most of the uncaptured observations have large, negative growth rate values. More specifically, only 0.55 on the observations with negative growth values fell within the training bounds. This suggests that the model was not properly calibrated for extreme market declines. This concern is reinforced by Figure 15(b), which shows that very few observations from the test set fall within a two-standard deviation confidence ellipsoid on a cluster created from training set data.

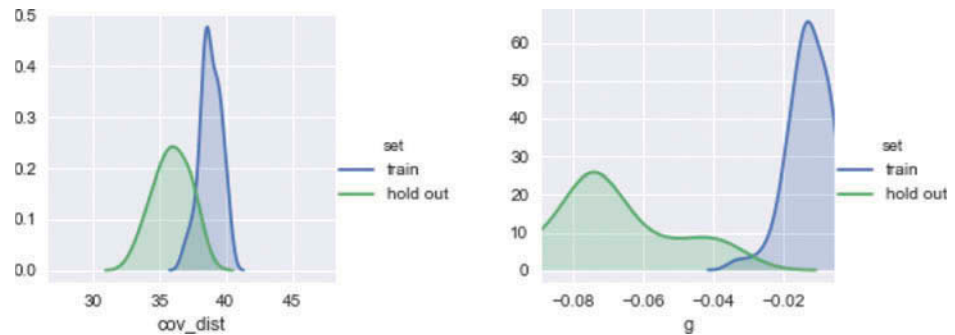
The cluster-specific feature values between training and testing sets did not show the same similarities seen in case study 1. Given the results shown in Figure 15(b), few observations were close to, or fell within the confidence ellipsoid of, the nearest cluster. This concern for the robustness of the characteristics of states was supported by the K-S test results. The statistical test conclusively rejected (significance level of 0.10) six of the eight features dismissing the idea data came from the same distribution. Figure 16(a and b) illustrate this result. Tests of these features resulted in p-values of approximately 0 in both cases.

Figure 15. Feature values and cluster assignment confidence.



(a) Feature values for Case 2 observations from training and test set. (b) Test set observation and cluster confidence ellipsoids.

Figure 16. Kolmogorov–Smirnov (K-S) test: feature comparison.



(a) The covariance distance for training and hold out testing set for Cluster 3. Statistically different. (b) The growth vector values for training and hold out testing set for Cluster 0. Statistically different.

Specifically for the growth vector in Figure 16(b), the values from the holdout set are skewed left indicating that there were a higher concentration of large, negative values observed in the test set than in the training set. The remaining two features—both from the same, sparsely populated cluster in the test set (Cluster 1)—did not have sufficient observations to conduct a test.

The test for the Markovian property yielded mixed and inconclusive results. The sample size and distribution of observation in the test set clusters did not provide enough expected observations for many of pairwise combinations to be statistically tested. Only 3 of 16 pairwise combinations met the recommended minimum sample size needed to perform the test; however, all those combinations showed statistical evidence of the Markov property. These combinations fell along the diagonal of the transition probability matrix representing the self-transitions for Clusters 1, 2 and 3. The p-values associated with the test of Cluster 1, 2, and 3 were 0.61, 0.92, and 0.72, respectively, presenting very little evidence in support of rejecting the hypothesis that the observations came from different distributions and therefore are not Markovian.

The next test evaluates whether the transition probability matrices between the training and test sets shown in Figure 17(a and b) are statistically similar. The χ^2 Test results conclusively reject the hypothesis with a p-value of approximately 0.

This case study presents an example of a clustering schema that did not produce a robust model of market states and the metrics from which the conclusion about model quality could be drawn. The training set was not representative of data from the holdout, e.g. test set, which could be observed by the percentage of testing set data values that fell within the bounds of the training set in addition to the number of testing set observations that were encompassed within the two-standard deviation confidence ellipsoids. The K-S test demonstrated that feature values were not distributed in a comparable manner. The most populated states (1, 2, 3) showed signs of the Markovian property with respect to self-transitions. However, the transition probability matrices comparison between sets showed that similar transitional behavior did not

Figure 17. Comparison of transition probability matrices.

$$P_{train} = \begin{pmatrix} 0.70 & 0.11 & 0.11 & 0.08 \\ 0.09 & 0.65 & 0.24 & 0.02 \\ 0.08 & 0.12 & 0.66 & 0.11 \\ 0.16 & 0.02 & 0.26 & 0.55 \end{pmatrix}$$

(a) Training set probability matrix

$$P_{test} = \begin{pmatrix} 0.94 & 0.00 & 0.00 & 0.06 \\ 0.00 & 0.00 & 1.00 & 0.00 \\ 0.05 & 0.00 & 0.91 & 0.05 \\ 0.00 & 0.25 & 0.50 & 0.25 \end{pmatrix}$$

(b) Testing set probability matrix

persist, and the solution presented was not a robust one. If this is found to be the case for a given clustering schema then the training period and number of clusters should be reconsidered to find a configuration that produces results that are similar enough to meet the user's specifications.

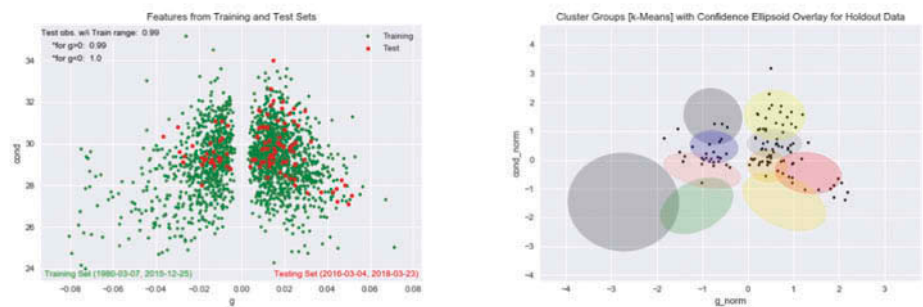
5.4. Case study 3: 1980 to present

The previous example illustrated the risks of using a state-based approach when many observations from the model's evaluation set fall outside the range of observations used to create the model. The Case 2 scenario results in model states that do not adequately capture the behavior of the period being evaluated. A solution to this problem is to examine more observations by expanding the training set used to define the state-based model. By including more observations, a more complex and granular model can be created. Case 3 examines a larger subset of data than the other two case studies. As shown in Figure 18(a) the model is created using data from 1980 through 2015 and then evaluated on 2016 to 2017 data. Further examination of the observations finds that all observations from the testing set fall within the bounds of the training data. Given the size of the data set, a more complex 10-state model is created as shown in Figure 18(b). Note that the confidence ellipsoid for each state contains a considerable number of observations on which to derive the state centroid.

Figure 18(b) shows observations from the test set and an overlay of confidence ellipsoids for the states defined by the training set data. Some ellipsoids do not have test set observations fall within their bounds. This is not unexpected given that the test set is comprised of data from a two-year period of time and instances of all 10 states from the model may not occur during that period of time. Conversely, some observations fell within regions where multiple ellipsoids overlapped suggesting membership in multiple states. Approaches for addressing such situations are left to the user based on preference and the intended use of the information derived from the state-based analysis; however, a future paper presents techniques for addressing such cases in practical applications, such as techniques for incorporating the probability of state membership into decision-making.

Increasing the number of states resulted in more pairwise combinations of states that could be tested for the existence of the Markovian property in the system and also resulted in higher ratio of

Figure 18. Feature values and cluster assignment confidence.



(a) Feature values for all case study observations from training and test set. (b) Test set observation and cluster confidence ellipsoids.

Table 4. Summary of Markov Property tests for all case studies

Case Study	Eligible Pairs	Rej. _{$\alpha=0.05$}	Rej. _{$\alpha=0.10$}
1	8	2	3
2	3	0	0
3	18	1	3

tests that failed to reject that the property existed. Of the possible 100 pairwise combinations, 18 were able to be tested. Of these 18 combinations, only 1 was rejected at a significance level of 0.05 and an additional 2 were rejected at a level of 0.10. This is higher concentration of combinations that exhibited evidence of the Markov property than the example presented in Case Study 1. While Case Study 2 presented an example in which all combinations showed evidence of the property, the small number of combinations eligible for testing failed to produce conclusive results. These results discussed are summarized in Table 4. The test results illustrate the benefit of defining the markets in a more granular and detailed manner by showing the persistence of transitional behavior between states by demonstrating the existence of the Markov property.

This example shows the value in increasing the level of granularity and detail in the states by increasing the number of states; however, the optimal number of states, k , is dependent of the model and data set in question. The process of determining the number of states is addressed in a future paper.

6. Conclusions

This article presents a new methodology for classifying periods of time in the financial markets. A discrete, state-based approach is used to subdivide time. State membership determinations are made using machine learning algorithms, which make assignments based on a two-dimensional feature set of expected growth and covariance values. The selection of features on which to make clustering decisions comes from traditional finance theory, specifically the research of Markowitz (1952). States are then defined without a preconception or bias as to which characteristics should be possessed by any state. Instead, states are created by clustering together observations with similar features.

The article goes on to propose an approach for making probabilistic determinations of state membership for all observations in a holdout set based on proximity to the centroid of each state created from the training data. The numbers of clusters selected for the model are crucial, and therefore future research into this selection process is needed, given the importance of the decision. It is imperative that the future research determines an appropriate metric on which to make this decision based on the objective of the clustering applications. Accounting for the application of the clustering methodology plays an important part in the determination of the metrics used.

The article discusses the importance of evaluating the efficacy of the states produced through this approach and presents metrics for examining the behavior and similarities between the base states and closest states in a testing set. Consistencies in state transition dynamics can be examined using a Markov model framework and then examine the transition behavior statistically using the Chi-Squared test. An application of the Kolmogorov–Smirnov (K-S) test is presented to examine the persistence of feature values between states in the training and testing data sets.

Future research could build on the foundation of work presented within this article in several areas. First, the development of a process to select the number of states for the model. A two-state model may not provide enough resolution to make well-informed decisions. A formalized methodology should be researched and developed that can provide a framework for making this determination given the data and intended use of the model. Second, research should be conducted into the practical applications of this underlying methodology. One practical application, in particular, recommended for future research is portfolio design, and the ways to incorporate this methodology into that pursuit. Several specific research questions related to portfolio design include examining the value of using separate data sets for creating states and then for making asset allocation decisions, how to make portfolio allocation decisions given state membership information, and how to make asset allocation decisions using probabilistic state membership information. Lastly, a richer feature set should be researched to gain additional insight into the quality of states created by incorporating additional features. Potential features for examination include higher order moments, such as skewness or kurtosis of the price data.

Funding

The authors received no direct funding for this research.

Author details

Matthew W. Burkett¹

E-mail: mburky@gmail.com

William T. Scherer¹

E-mail: wts@virginia.edu

Andrew Todd¹

E-mail: aetodd@gmail.com

¹ Systems and Information Engineering, University of Virginia, Charlottesville, USA.

Citation information

Cite this article as: Are financial market states recurrent and persistent?, Matthew W. Burkett, William T. Scherer & Andrew Todd, *Cogent Economics & Finance* (2019), 7: 1622171.

Notes

1. The source of the data is the Federal Reserve Economic Data (FRED) <https://fred.stlouisfed.org>.
2. The first half and second half of 2008 represent distinctly different periods of time in financial market history. Following the decline of global equity markets in Q4 2007, the many markets did not experience large gain or losses in the first half (Q1, Q2) of 2008; however, in the second half (Q3, Q4), markets worldwide experienced dramatic losses.
3. The centroid is a way to represent the point at the center of a cluster. It is a measure of central tendency of each feature for the members of a cluster. It can be derived by calculating the arithmetic mean of each feature (Kevin, 2012).
4. The steady-state probability of a Markov chain represents the long-term probability of being in a state at any point in time. It can also be thought of as the percentage of time spent in each state in the long term. Given that the steady-state is based on a long-term view of the system, the steady-state values are independent of the initial starting state (Bertsekas, 2002).
5. Seven assets were included—four equity indexes and three bonds products. The domestic equity indexes included were: S& P 500, NASDAQ Composite, Dow Jones Industrial Average, and the Russell 2000. The bond products included were: 10-year Treasury, Moody's AAA, and Moody's BAA.

References

- Badge, J. (2012). Forecasting of Indian stock market by effective macro-economic factors and stochastic model. *Journal of Statistical and Econometric Methods*, 1(2), 39–51.
- Bailey, D. H., Borwein, J., Lopez de Prado, M., & Zhu, Q. J. (2016). The probability of backtest overfitting. *Journal of Computational Finance*.
- Bartlett, M. S. (1951). "The frequency goodness of fit test for probability chains." In *Mathematical Proceedings of the Cambridge Philosophical Society*, 47, pp. 86–95, Cambridge University Press.
- Bertsekas, D. P. (2002). *Introduction to probability: Dimitri P. Bertsekas and John N. Tsitsiklis*. Athena Scientific. CA: Belmont.
- Billingsley, P. (1961). Statistical methods in Markov chains. In *The annals of mathematical statistics*.
- Devore, J. (2015). *Probability and statistics for engineering and the sciences*. Cengage Learning. CA: Belmont.
- Draper, N. R., & Smith, H. (2014). *Applied regression analysis*. New York, NY: John Wiley & Sons.
- Farmer, J. D., Gallegati, M., Hommes, C., Kirman, A., Ormerod, P., Cincotti, S., ... Helbing, D. (2012). A complex systems approach to constructing better models for managing financial markets and the economy. *The European Physical Journal Special Topics*, 214(1), 295–324. doi:10.1140/epjst/e2012-01696-9
- Forstner, W., & Moonen, B. (1999). A metric for covariance matrices.
- Friendly, M., Monette, G., & Fox, J. (2013). Elliptical insights: Understanding statistical methods through elliptical geometry. *Statistical Science*, 28(1), 1–39. doi:10.1214/12-STS402
- Ghezzi, L. L., & Piccardi, C. (2003). Stock valuation along a markov chain. *Applied Mathematics and Computation*, 141(2–3), 385–393. doi:10.1016/S0096-3003(02)00263-1
- Guidolin, M., & Timmermann, A. (2007). Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control*, 31(11), 3503–3544. doi:10.1016/j.jedc.2006.12.004
- Hamilton, J. D. (1989, Mar). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57, 357. doi:10.2307/1912559
- Hazewinkel, M. (2001). Kolmogorov-Smirnov test. In *Encyclopedia of mathematics* (pp. 978). Springer, ISBN.
- Johnson, N. L., & Kotz, S. (1970). *Discrete distributions: Continuous univariate distributions-2*. New York, NY: Houghton Mifflin.
- Kevin, P. M. (2012). *Machine learning: A probabilistic perspective*. Cambridge, UK: MIT Press.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.
- Neftci, S. N. (1984). Are economic time series asymmetric over the business cycle? *Journal of Political Economy*, 92(2), 307–328. doi:10.1086/261226
- Owyang, M. T., & Ramey, G. (2004). Regime switching and monetary policy measurement. *Journal of Monetary Economics*, 51(8), 1577–1597. doi:10.1016/j.jmoneco.2004.01.004
- Romesburg, C. (2004). *Cluster analysis for researchers*. Lulu. com. CA: Belmont.
- Scherer, W. T., Adams, S., & Beling, P. A. (2018). On the practical art of state definitions for Markov decision process construction. *IEEE Access*, 6, 21115–21128. doi:10.1109/ACCESS.2018.2819940
- Scherer, W. T., & Glagola, D. M. (1994). Markovian models for bridge maintenance management. *Journal of Transportation Engineering*, 120(1), 37–51. doi:10.1061/(ASCE)0733-947X(1994)120:1(37)
- Whitelaw, R. F. (2000). Stock market risk and return: An equilibrium approach. *The Review of Financial Studies*, 13(3), 521–547. doi:10.1093/rfs/13.3.521



© 2019 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format.

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



***Cogent Economics & Finance* (ISSN: 2332-2039) is published by Cogent OA, part of Taylor & Francis Group.**

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

