



Intertemporal Similarity of Economic Time Series: An Application of Dynamic Time Warping

Philip Hans Franses¹ · Thomas Wiemann¹

Accepted: 14 April 2020
© The Author(s) 2020

Abstract

This paper adapts the non-parametric dynamic time warping (DTW) technique in an application to examine the temporal alignment and similarity across economic time series. DTW has important advantages over existing measures in economics as it alleviates concerns regarding a pre-defined fixed temporal alignment of series. For example, in contrast to current methods, DTW can capture alternations between leading and lagging relationships of series. We illustrate DTW in a study of US states' business cycles around the Great Recession, and find considerable evidence that temporal alignments across states dynamic. Through cluster analysis, we further document state-varying recoveries from the recession.

Keywords Business cycles · Non-parametric method · Dynamic time warping

JEL Classification C14 · C50 · C55 · C87 · E32

1 Introduction

In this paper we illustrate the non-parametric technique Dynamic Time Warping (DTW) to examine similarities of business cycles. DTW is a widely applied algorithm in non-economic fields such as speech, pattern and movement recognition [see, for example, Cedras and Shah (1995), Geiger et al. (1995), and Sakoe and Chiba (1978)] and data mining [see, for example, Keogh and Pazzani (2000), among others]. While its non-parametric similarity measure has been shown to be superior to other measures such as the Pearson correlation coefficient (Petitjean et al. 2011), economic research has not yet used Dynamic Time Warping to its potential. To the best of our knowledge, Wang et al. (2012) and Raihan (2017) are the exceptions, where the DTW technique is used in its basic standard format. In our study,

✉ Philip Hans Franses
franses@ese.eur.nl

¹ Econometric Institute, Erasmus School of Economics, Burgemeester Oudlaan 50, 3062PA Rotterdam, The Netherlands

we design the DTW technique in particular for our purposes to identify dynamic leading and lagging relationships between time series. We illustrate the application of our modified DTW technique by investigating business cycle similarities across US states.

There is substantial literature on similarities across time series variables, predominantly using parametric techniques. Usually, it is assumed that when one variable is leading the other, it does so throughout the entire sample. Bai and Ng (2004) consider a dynamic factor model with a common factor structure. Engle and Kozicki (1993) examine common cyclical and other features. Cubadda et al. (2009) study common autoregressive parts across individual series as an indication of common features in multivariate models. Fröhwrth-Schnatter and Kaufmann (2008) propose a clustering method based on mixture models, where each mixture component follows the same underlying model specification. Due to the parametric nature of these approaches to time series' similarity, however, neither can adequately describe a time-varying dynamic specification. In particular, and to put it simply, these methods restrict that variables y_t and x_t are linked via, for example, the (vastly simplified) specification $y_t = \alpha + \beta x_{t-j} + \varepsilon_t$ with j fixed over the sample. In this paper we propose to study co-movements using the DTW technique where we allow for the possibility that j is time-varying and may switch from positive to negative or vice versa.

A prominent strand of economic literature for which this seems particularly relevant is the research on the similarity of business cycles. Here, most studies analyse correlation structures with, for example, common factor models or their adaptations (Francis et al. 2017). Recently, also non-parametric clustering techniques have received increased attention. For example, Crone (2005) and Papageorgiou et al. (2017) consider K-means clustering to group business cycles of US states and European countries, respectively. A common caveat among these previously considered approaches, however, is the strict assumption of identically timed business cycles. As suggested by the findings of Hamilton and Owyang (2012), this is potentially problematic as a substantial amount of idiosyncratic business cycles seems to stem from differential timing around national recessions. If one is interested not only in the co-movement at a particular point in time, but also the similarity in the extent or gravity of the economic movements, existing techniques only allow for limited insights if their results are largely driven by even slight temporal shifts of time series. DTW can alleviate such concerns by directly incorporating potentially dynamic timing differences in the analysis.

For a first and intuitive impression on how DTW can contribute to existing literature, consider Fig. 1. The graph shows the DTW-produced alignment of quarterly, seasonally adjusted real GDP of Florida and Idaho for the period 2006Q1 to 2017Q4. The lines that link two points of the respective series show their closest correspondence. While in some cases the connected points concern the same quarter, the figure indicates that in most cases, earlier or later quarters are apparently a better match. This feature is the key reason that we rely on this non-parametric

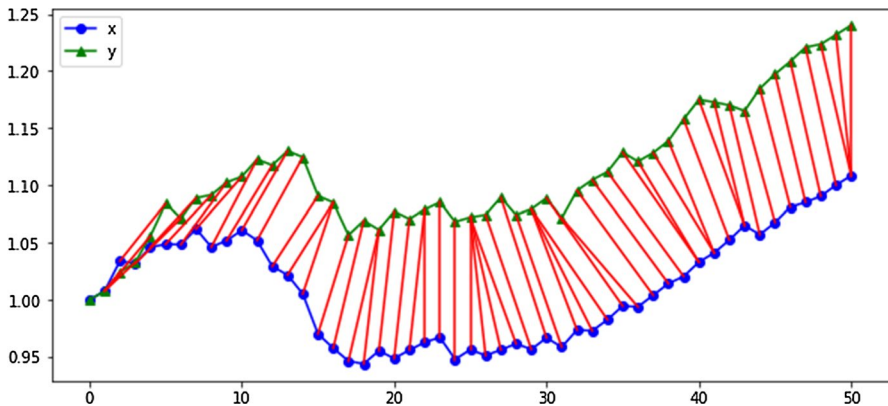


Fig. 1 Temporal point mapping. Quarterly real GDP of Florida (x , blue) and Idaho (y , green) for the period 2006Q1 to 2017Q4, standardized at 1.00 for 2006Q1. (Color figure online)

technique.¹ From Fig. 1 we see that the time-alignment of the two series, denoted by j earlier, seems to vary over the sample, ranging over values such as 0, 1, -1 and even -4 and 4. One could now decide to make j a time-varying parameter, for example j_t , however, given the size of typical samples on real GDP data, a non-parametric technique without the need for high degrees of freedom is more appropriate. Dynamic Time Warping, at least in the variant proposed below, seems to be a useful technique: it not only illustrates the link of two time series visually but it also provides an estimate of their temporal alignment, as we will show below.

The introduction of DTW—as a computer science method originally developed for pattern recognition—complements the recent developments in machine learning applications in economics, although with a clearly distinct contribution. In reviews of the relevant literature, Athey (2017) and Mullainathan and Spiess (2017) illustrate the potential of machine learning for advanced methods on causal inference and for predictive applications in economics. For DTW, we suggest that it contributes most to descriptive data analysis by revealing temporal dynamics between time series and allowing for assessment of their similarity. In a similar manner as machine learning methods can complement conventional econometric models, we believe DTW can complement—and in some contexts replace—the conventional Pearson correlation coefficient and other similarity measures. At the same time, the potential of this method is beyond its descriptive abilities. For example, Raihan (2017) makes promising suggestions for the use of DTW techniques for predictive purposes.

In the rest of this paper, we proceed as follows. In Sect. 2, we outline standard DTW and introduce our modifications for our analysis of economic time series. Our adjusted DTW is illustrated by an application to real GDP series of three US states.

¹ Funashima (2017) uses wavelet analysis for approximately similar reasons. Alternatively, one can consider time varying spectral analysis, see Sejdic et al. (2009) for an overview of various methods. An important aspect of frequency analysis is that one needs a large enough sample to run the various, sometimes complicated, computations. Our DTW technique is much simpler, and it also works for small samples like the one considered in the current paper.

In Sect. 3, we discuss an application of DTW-based time series clustering to assess similarities of all mainland US state real GDP series. Section 4 concludes with some ideas for further research.

2 Dynamic Time Warping

This section outlines standard DTW and introduces our modification for our analysis of business cycle similarities. Throughout, we focus on the mathematical specification of the measure and its economic interpretation. For calculation of the DTW alignment by means of dynamic programming, Müller (2007) provides an excellent description of the algorithm.

DTW defines a distance-minimising temporal alignment between two time series $X = [x_1, x_2, \dots, x_N]$ and $Y = [y_1, y_2, \dots, y_M]$. For this purpose, we first calculate the $N \times M$ distance matrix D , where the (i, j) th element of the matrix is given by the distance between the points x_i and y_j . Each element of the matrix thus corresponds to a mapping of two points between the time series. The optimal alignment is then given by the *warping path* p that minimises the cumulative distance of all mapped point-pairs on the path, denoted by

$$\gamma(i, j) = d(x_i, y_j) + \min[\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)] \quad (1)$$

with $d(x_i, y_j)$ being a distance measure between points x_i and y_j .

The DTW problem can be solved straightforwardly by means of an $O(MN)$ dynamic programming algorithm as given by, for example, Müller (2007). The resulting minimal *warping cost*, $\gamma(N, M)$, constitutes a non-parametric similarity measure of time series. When aiming to compare time series of differing length, it seems useful to assess the warping cost per considered time series element. In this case, we can consider

$$\Gamma(N, M) = \frac{\gamma(N, M)}{M + N} \quad (2)$$

where N and M are the lengths of the considered time series.

2.1 The Growth-Based Distance Function

While any function $d(x_i, y_j) \rightarrow \mathbb{R} \geq 0$ could be considered as a distance measure, in standard DTW it is typically defined as either $|x_i - y_j|$ or $(x_i - y_j)^2$ [see, for example, Raihan (2017) and Wang et al. (2012)]. In the context of comparing economic time series such as business cycles, however, this choice does not seem perfectly appropriate. For example, it would ignore the definition of expansion and contraction periods if the value of only one period at a time is taken into account. More ideally, one therefore also considers an observation's relationship to its earlier and subsequent neighbours to identify the features of the time series.

For these reasons, we consider the *feature*-based distance function of Xie and Wiltgen (2010). It constitutes a well suited measure that captures both the overall shape of the time series and the local trend around the points. The authors illustrate its advantage over standard, *value*-based DTW and the widely applied derivative DTW of Keogh and Pazzani (2001) by means of a simulation study.

We will adjust the feature-based distance function to avoid issues such as substantial differences in magnitudes of levels. The local feature of a particular observation is originally defined by a 2-element vector that summarises the slope to its left- and right-hand side neighbour. Rather than the absolute slope, we consider the growth rate between the periods to capture local trends. This modification fits common business cycle analysis. The local feature of observation i of a series X is then given by

$$f_{local}(x_i) = \left[\frac{x_i - x_{i-1}}{x_{i-1}}, \frac{x_{i+1} - x_i}{x_i} \right] \tag{3}$$

The global feature is adjusted in a similar manner. Instead of the absolute deviation of a series' point to its left- and right-hand side as proposed by Xie and Wiltgen (2010), we calculate its relative deviation to capture the shape of the particular series while correcting for substantial differences in magnitude across the set of time series. The global feature of observation i of a series X is then given by

$$f_{global}(x_i) = \left[\frac{x_i - \sum_{k=1}^{i-1} \frac{x_k}{i-1}}{\sum_{k=1}^{i-1} \frac{x_k}{i-1}}, \frac{x_i - \sum_{k=i+1}^N \frac{x_k}{N-i}}{\sum_{k=i+1}^N \frac{x_k}{N-i}} \right] \tag{4}$$

Finally, our modified feature-based distance function is thus given by

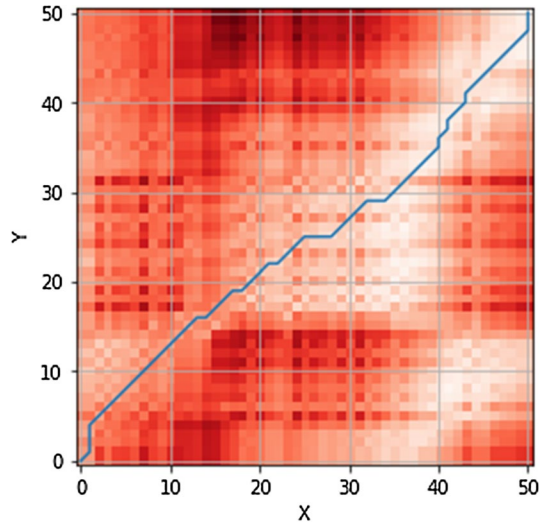
$$\begin{aligned} d(x_i, y_j) &= d_{local}(x_i, y_j) + d_{global}(x_i, y_j) \\ &= |f_{local}(x_i)_1 - f_{local}(y_j)_1| + |f_{local}(x_i)_2 - f_{local}(y_j)_2| \\ &\quad + |f_{global}(x_i)_1 - f_{global}(y_j)_1| + |f_{global}(x_i)_2 - f_{global}(y_j)_2| \end{aligned} \tag{5}$$

where the subscripts 1 and 2 refer to the corresponding elements in the vector. In short, our choice in (6) allows the data to have trends, and this is relevant for many economic time series. By balancing the local and global features, we can compare the data points as such, and evaluate them relative to their adjacent observations as well as to the overall pattern in the data.

3 Identification of Temporal Dynamics

Before we move to the analysis of all US states, we first consider Florida, Idaho and California and consider quarterly real GDP of US states from 2006Q1 till 2017Q4 obtained from the Bureau of Economic Analysis (BEA). For now, we focus on the insights obtained through DTW on the leading and lagging

Fig. 2 Warping Path and Distance Matrix. Quarterly real GDP of Florida (x) and Idaho (y) for the period 2006Q1 to 2017Q4. Regions of low and high cost (distance) are indicated by light and dark colours, respectively



relationship between two time series. A discussion of the comparison of time series based on their similarities will follow in Sect. 4.

First, let us consider the example of Florida's and Idaho's GDP series presented in Fig. 1 of the introduction. When leads and lags would be constant throughout the sample, the lines connecting the two time series would all have a similar angle relative to one of the two variables. From the slope of the alignments in Fig. 1, we can however see that Idaho (green) leads Florida (blue) in the first half of the sample as the connecting lines face the north-east, so to say, while around observations 27 to 30, the connecting lines face the north-west. From around halfway the sample onward, the dynamics of the two series seem to change and Florida's GDP growth leads Idaho's for the remainder of the sample.

An alternative visualisation of the DTW results is given in Fig. 2. The distance-minimising mapping of points, shown in red in Fig. 1, is given by the warping path in blue. It is plotted on the distance matrix that was computed using our distance measure introduced above. A non-shifted temporal alignment between the two series would be indicated by a fully diagonal warping path, whereas deviations from the diagonal indicates leading and lagging relationships of the series. Straightforwardly, we can thus infer the lead of Idaho over Florida for the first half of the sample by the left-hand deviations from the diagonal. Similarly, we can observe the change in the relationship at around halfway the sample, from which onward Florida leads Idaho by the right-hand deviations from the diagonal. These findings seem to suggest that Idaho's GDP was affected by the great recession in 2008/2009 sooner than Florida (a lead of Idaho), but Florida's GDP began to recover earlier (a lead of Florida). DTW therefore identifies substantial time-varying temporal shifts between the time series. As the conventional correlation coefficient solely considers a pre-specified, fixed temporal alignment of time series, these results highlight the advantage of DTW when interested in the similarity of series beyond a temporal alignment.

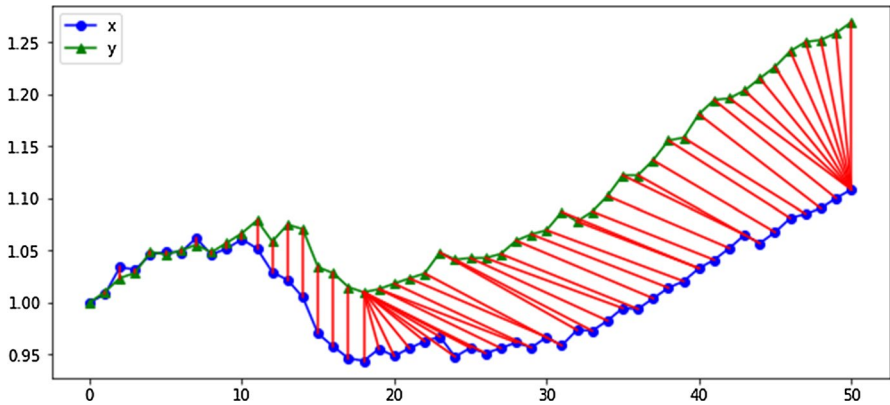
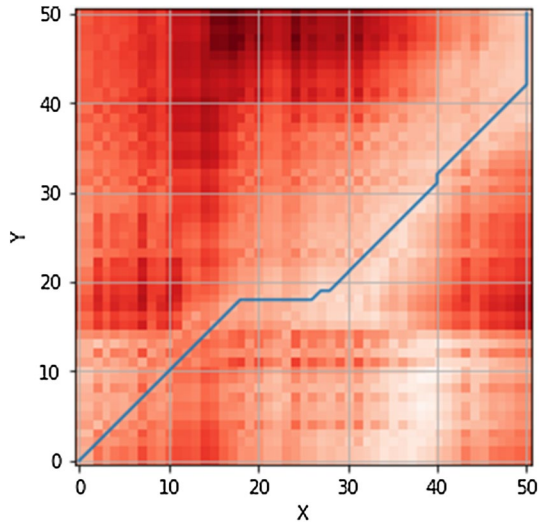


Fig. 3 Temporal Point Mapping. Quarterly real GDP of Florida (x, blue) and California (y, green) for the period 2006Q1 to 2017Q4. (Color figure online)

Fig. 4 Warping path. Quarterly real GDP of Florida (x) and California (y) for the period 2006Q1 to 2017Q4



To provide another, more pronounced example, we also analyse Florida’s and California’s real GDP series by means of DTW. Figure 3 presents the two series and their point mapping. Figure 4 shows the corresponding distance matrix and the warping path. As indicated by the vertical point-mappings and the diagonal warping path, the two GDP series are not shifted for the first four and a half years. After that, we see a substantial change in the relationship, where California leads Florida’s GDP series for the remainder of the sample. This is shown by both the non-vertical point-mappings and the right-hand deviation from the diagonal of the warping path in Fig. 4. These graphs suggest that both Florida and California were affected by the recent great recession at about the same time, but California’s GDP seemed to have recovered sooner than Florida’s (that is, there is a lead of California).

4 Time Series Clustering

Apart from insights on the pairwise leading and lagging relationships between two series, it is also of interest to assess the similarity of time series. When only a small set of series is under investigation, we could simply compare the minimal warping cost per observations as introduced in (2). Concerning the already discussed examples, we find that the warping cost for Florida and Idaho is 0.036, and 0.048 for Florida and California. Florida is thus more similar to Idaho than to California as assessed by DTW.

In many cases, however, the set of time series is so large that such a comparison becomes bothersome if not infeasible. For G time series, one would have to assess $\frac{G(G-1)}{2}$ warping distances. As an alternative, time series clustering provides the possibility to make structured inferences about the similarity of large sets of time series in a concise manner. This is achieved by organizing the series into homogeneous groups, minimizing intra-group dissimilarity and maximizing inter-group dissimilarity. Time series clustering based on DTW would allow to discover data structures and it has proven to be a useful method, for example, for the purpose of data mining (Aghabozorgi et al. 2015; Liao 2005). In the following, we aim to illustrate DTW's contribution in this regard also in our context of investigating similarity structures across US states' real GDP series.

4.1 DTW-Based K-Means

A straightforward clustering approach is the K-means algorithm originally developed by MacQueen et al. (1967). It aims to minimize the total difference of the cluster's sequences to the respective cluster average by iteratively reassigning sequences to clusters whose average is most similar to the sequence and recalculating the cluster averages in each iteration. The algorithm converges when the total intra-cluster differences cannot be reduced further. In the below, we briefly illustrate how K-means can be adjusted to reflect time series similarity as captured by growth rate-based dynamic time warping.²

The K-means algorithm requires a well-defined 'average' of a set of series. This is not trivial when considering DTW distance measures. Petitjean et al. (2011) address this issue by introducing a global averaging method, DTW Barycenter Averaging (DBA). It consists of a heuristic strategy to approximate an average that accounts for temporal shifts in the spirit of DTW. DBA iteratively refines an initial average sequence to minimize its squared warping distance to the set of to-be-averaged series. This is achieved by calculating the temporal alignment of each series to the currently best average, and then defining the mean of all time series points that are temporally aligned with a particular point of the average as its updated version. Note that this allows the DTW average—or DBA—to reflect potentially dynamic

² An excellent illustration of conventional K-means is provided by, for example, Liao (2005a, b) or Aghabozorgi et al. (2015). For an application of (non-DTW) K-means in an economic context, see, for example, Sullivan and Miller (1996).

temporal shifts across time series. In particular, a specific point of the average might be associated with multiple points of a time series.³ To adapt this method to the application of business cycles, we consider the average of associated growth rates rather than each point's absolute value. The DTW-based average time series can easily be constructed by iterative multiplication of the growth rates and an initial index value (for example 1).

As K-means might converge to local optima, it has become common practice in economic research to perform multiple instances with random starting clusters for the search of a global maximum. To reduce the number of required random instances, we optimize the initial clustering similar to the idea of the DTW-based K-means++ algorithm as in Zhang and Hepner (2017). Rather than initializing clusters by dividing the set of time series entirely randomly, each of the series has a probability to be assigned as the initial centroid to one of the k clusters in disproportional to its similarity with existing centroids. In particular, one of the considered time series is chosen at random to be the centroid of the first cluster. For the remaining time series, we calculate the probability p_i is given by

$$p_i = \frac{\Gamma_i^2}{\sum_{j=1}^G \Gamma_j^2} \quad (6)$$

where $\Gamma_i = \sum_{c \in C} DTW(x_i, c)$ with C being the set of existing centroids, x_i being the i th time series, and DTW being the Dynamic Time Warping distance as defined before. A time series is selected as the centroid for the next cluster with probability p_i , and the step is repeated until all clusters are assigned to an initial centroid. The remaining time series are then grouped into the cluster with the 'closest' centroid as assessed by the DTW distances.

Starting at this initial clustering, K-means then proceeds by averaging the time series of each cluster by means of DBA and reassigning each time series to the cluster with the lowest DTW distance. This is continued until no series is reassigned and the procedure terminates.

We will now illustrate this procedure by analyzing US states' quarterly GDP around the period of the Great Recession.

4.2 An Application to US States

To illustrate the DTW-based clustering procedure, we proceed by analyzing US states' quarterly GDP for the period 2006Q1 till 2017Q4. Choosing the cluster amount⁴ of 8 we find the clusters as depicted in Fig. 5.

³ For an application in a non-economic context, see, for example, Gullo, Ponti, Tagarelli, Tradigo, and Veltri (2012).

⁴ The corresponding scree plot is in Fig. 15. No evident elbow can be identified, but with 9 or more clusters, more clusters become single-state-clusters. Note that this seems to be in line with Crone (2005), who found eight clusters for grouping contiguous US states with conventional K-means.

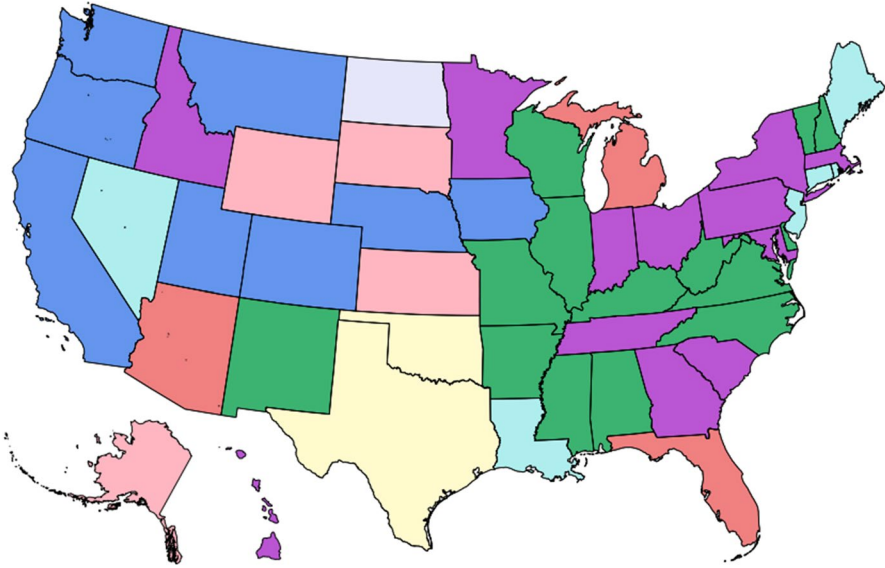


Fig. 5 Map of the US States with 8 clusters

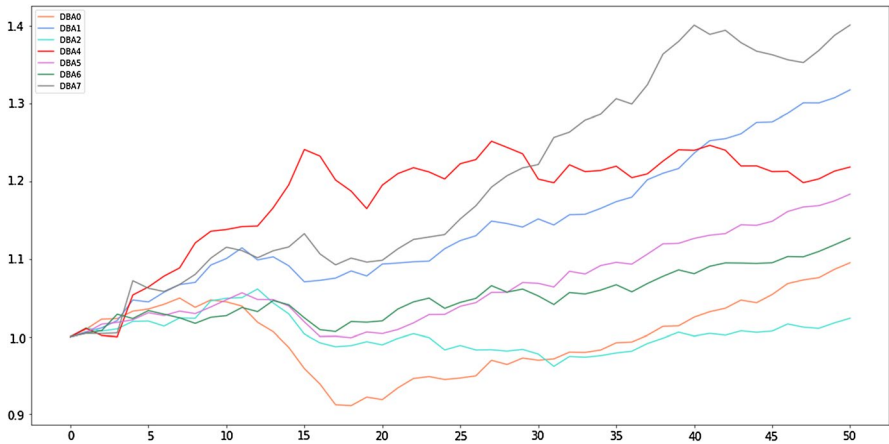


Fig. 6 Clusters’ DTW Barycenter averages. Only DBAs with more than one state are presented. (North Dakota is omitted.). The sample (horizontal axis) runs from 2006Q1 till 2017Q4 and the data are obtained from the Bureau of economic analysis (BEA)

To characterize the clusters, it is useful to consider their averages as computed by DBA. Figure 6 plots the averages of the clusters with more than one state. As DTW assesses similarity of series when allowing for temporal shifts between series, it does not seem particularly sensible to comment on the specific timing of certain features of the averages. In contrast, it is perfectly sensible to infer about their general shapes.

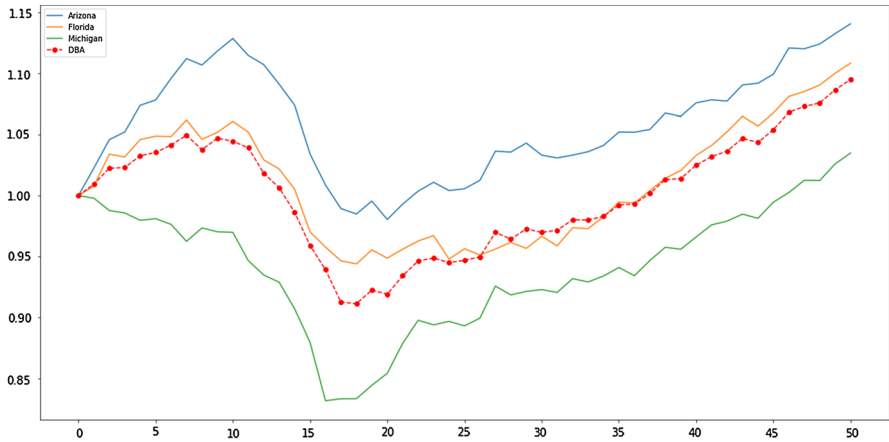


Fig. 7 Cluster 1: Arizona, Florida, Michigan. The sample (horizontal axis) runs from 2006Q1 till 2017Q4 and the data are obtained from the Bureau of Economic Analysis (BEA)

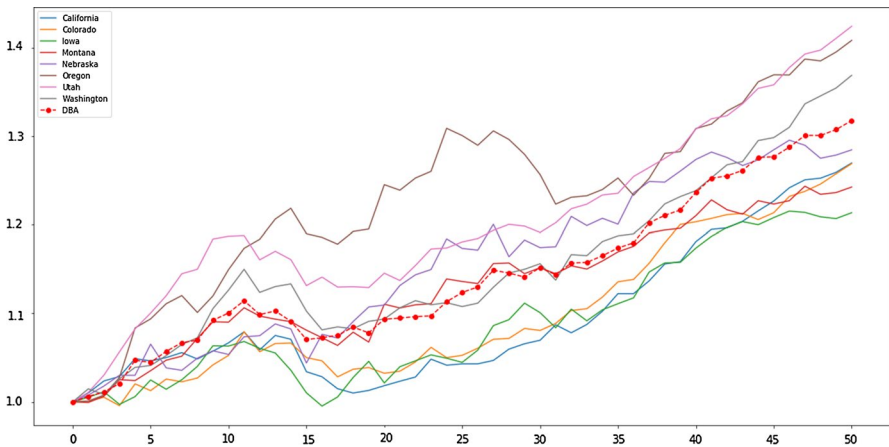


Fig. 8 Cluster 2: California, Colorado, Iowa, Montana, Nebraska, Oregon, Utah, Washington. The sample (horizontal axis) runs from 2006Q1 till 2017Q4 and the data are obtained from the Bureau of economic analysis (BEA)

If we zoom in on the graphs in Figs. 7, 8, 9, 10, 11, 12, 13 and 14, we see that there is much heterogeneity across the clusters. Figure 7 shows that Arizona, Florida and Michigan are in a cluster with states that suffered strongly from the financial crisis in 2008/2009, and it took many years for these states to recover. In stark contrast, for the states in Cluster 2, which are California, Colorado, Iowa, Montana, Nebraska, Oregon, Utah and Washington, it took less than 3 years to recover from the crisis, and after that they show substantial positive growth. These states came out very strong from the crisis. At the other end, the states in Cluster 3, which are Connecticut, Louisiana, Maine, Nevada, New Jersey and Rhode Island, did not

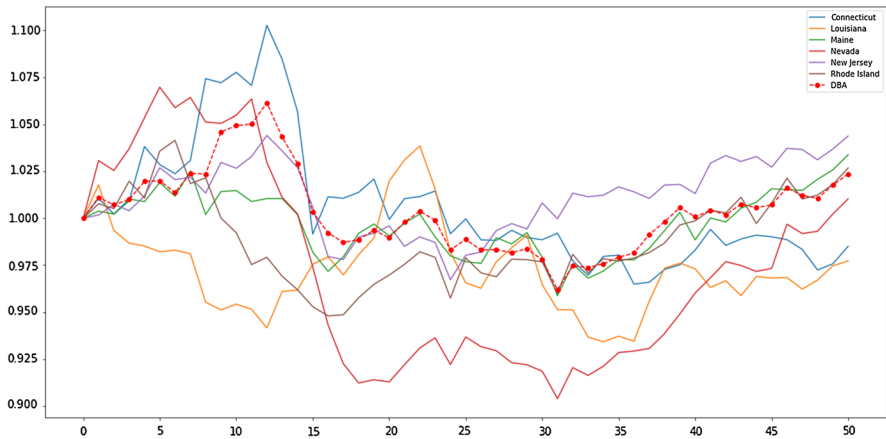


Fig. 9 Cluster 3: Connecticut, Louisiana, Maine, Nevada, New Jersey, Rhode Island. The sample (horizontal axis) runs from 2006Q1 till 2017Q4 and the data are obtained from the Bureau of economic analysis (BEA)

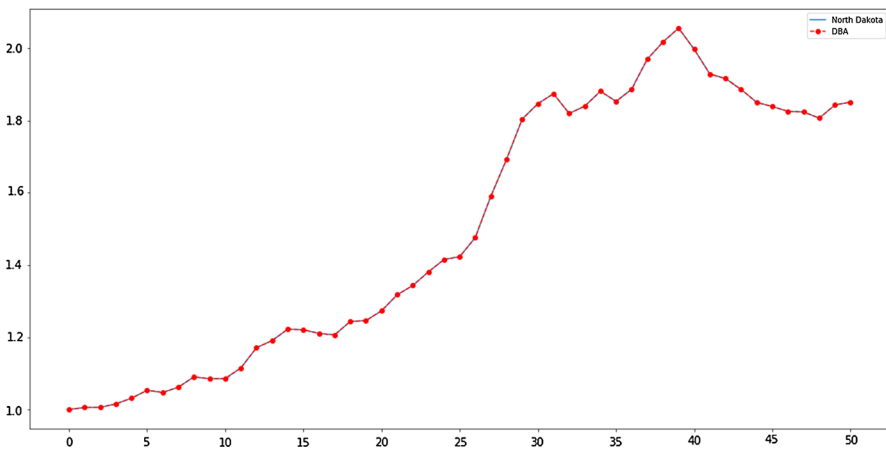


Fig. 10 Cluster 4: North Dakota. The sample (horizontal axis) runs from 2006Q1 till 2017Q4 and the data are obtained from the Bureau of economic analysis (BEA)

anymore attain the high levels of growth that they had before the crisis. Cluster 4 only contains North Dakota, and this state seems to follow its own idiosyncratic pattern. Alaska, Kansas, South Dakota and Wyoming in Cluster 5 did not seem to suffer from much setback due to the 2008/2009 crisis. Clusters 6 and 7 contain the largest contingents of states, for which the crisis caused just a year or two with a setback, after which growth continued with a steady pace. Finally, the states in Cluster 8, which are Oklahoma and Texas, seem not to have suffered at all from the crisis.

In sum, some clusters' DBAs indicate that some states did not have a downward impact from the recession at all. This seems to hold for the states in cluster 7, including



Fig. 11 Cluster 5: Alaska, Kansas, South Dakota, Wyoming. The sample (horizontal axis) runs from 2006Q1 till 2017Q4 and the data are obtained from the Bureau of economic analysis (BEA)

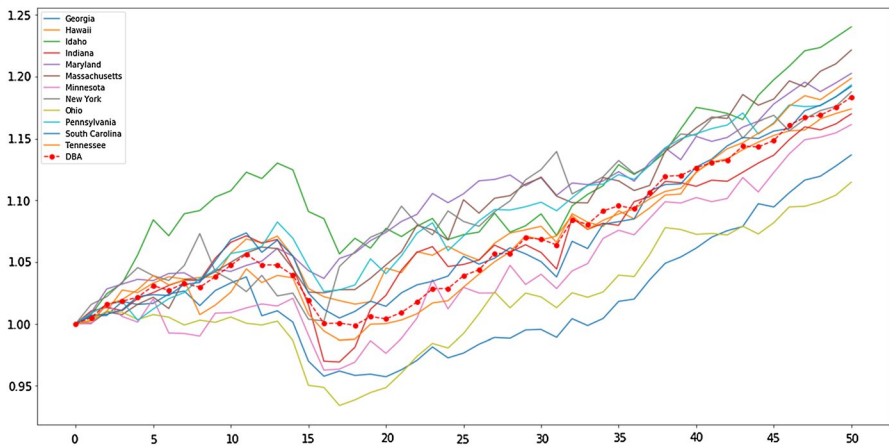


Fig. 12 Cluster 6: Georgia, Hawaii, Idaho, Indianan, Maryland, Massachusetts, Minnesota, New York, Ohio, Pennsylvania, South Carolina, Tennessee. The sample (horizontal axis) runs from 2006Q1 till 2017Q4 and the data are obtained from the Bureau of economic analysis (BEA)

states as Alabama, Kentucky, New Mexico and Wisconsin. In contrast, the states in cluster 3, which are Connecticut, Louisiana, Maine, Nevada, New Jersey and Rhode Island, did not seem to recover from the recession, even as late as by the end of 2017Q4.

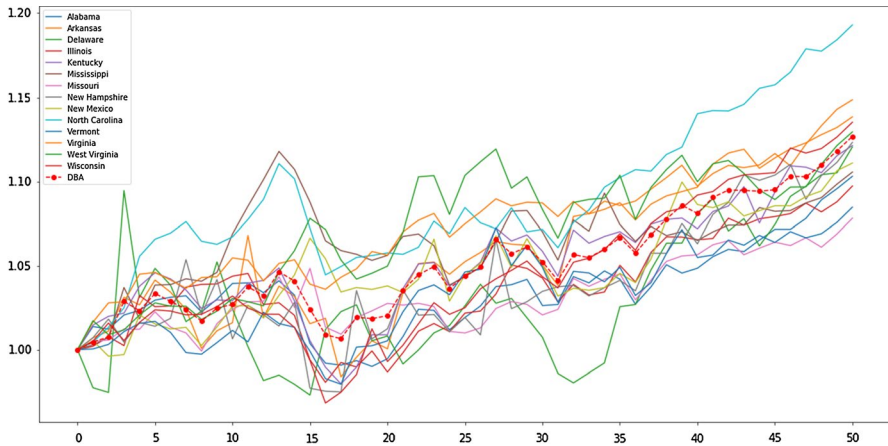


Fig. 13 Cluster 7: Alabama, Arkansas, Delaware, Illinois, Kentucky, Mississippi, Missouri, New Hampshire, New Mexico, North Carolina, Vermont, Virginia, West Virginia, Wisconsin. The sample (horizontal axis) runs from 2006Q1 till 2017Q4 and the data are obtained from the Bureau of economic analysis (BEA)

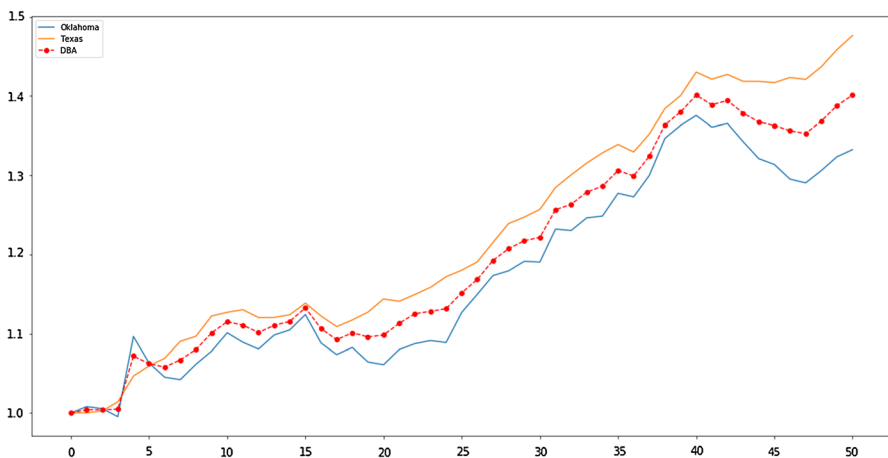


Fig. 14 Cluster 8: Oklahoma, Texas. The sample (horizontal axis) runs from 2006Q1 till 2017Q4 and the data are obtained from the Bureau of economic analysis (BEA)

5 Conclusion

We introduced the Dynamic Time Warping technique to examine potential similarities across business cycles, where we focused on quarterly real GDP. This non-parametric DTW technique is based on distances between observations, and seeks to find a minimum total distance as the optimal path. As such, it allows for the possibility that lead and lag relations can change during the sample period. With a few examples we showed that this switching behaviour is more common than rare. Parametric

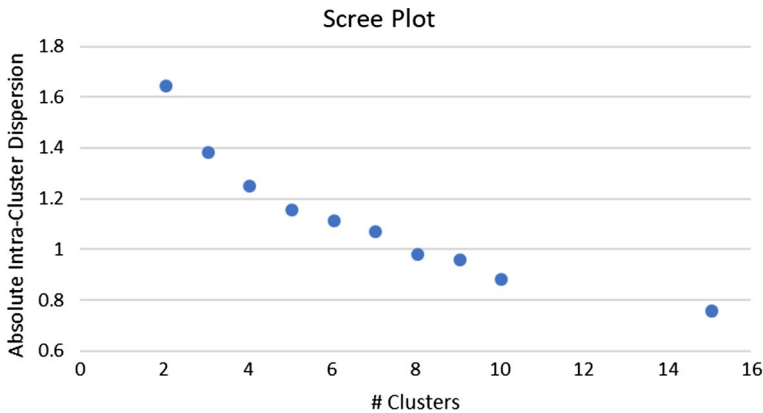


Fig. 15 Scree plot

models to allow for this behaviour will quickly become parameter-heavy, and given the commonly observed sample sizes, we believe that a non-parametric technique comes in handy.

The DTW technique is conceptually not complicated, although it may take some computer time to run the calculations, specifically when clusters need to be created. We envisage various other application areas for the DTW technique. In essence it is a method to connect moving objects, and as such many economic variables come to mind which may benefit from DTW analysis. Non-parametric count methods can be used to introduce statistical inference, to examine if a leading variable turned into a lagging variable after some event.

We proposed to use Dynamic Time Warping as a machine learning technique in economics and finance, here with an application to examine common patterns across economic variables. The study of common patterns is important, see also Corona et al. (2020), as it can lead to a better understanding of data features across multiple time series, and it can be beneficial for forecasting. With the advance of more and more detailed data, we foresee the study of common patterns as an important area for research. At the same time, when so many time series data become available, we also envisage a need for the application of machine learning techniques, see also Katris (2020), as these techniques can help to discover patterns that are not immediately obvious. More research into designing powerful machine learning methods to study common patterns is thus an important avenue for future research.

Acknowledgements We thank an anonymous reviewer for helpful comments.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission

directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aghabozorgi, S., Shirkhorshidi, A. S., & Wah, T. Y. (2015). Time-series clustering—A decade review. *Information Systems*, *53*, 16–38.
- Athey, S. (2017). The impact of machine learning on economics. *Economics of artificial intelligence*. Chicago: University of Chicago Press.
- Bai, J., & Ng, S. (2004). A PANIC attack on unit roots and cointegration. *Econometrica*, *72*(4), 1127–1177.
- Cedras, C., & Shah, M. (1995). Motion-based recognition a survey. *Image and Vision Computing*, *13*(2), 129–155.
- Corona, F., Poncela, P., & Esther, R. (2020). Estimating non-stationary common factors: Implications for risk sharing. *Computational Economics*, *55*(1), 37–60.
- Crone, T. M. (2005). An alternative definition of economic regions in the United States based on similarities in state business cycles. *Review of Economics and Statistics*, *87*(4), 617–626.
- Cubadda, G., Hecq, A., & Palm, F. C. (2009). Studying co-movements in large multivariate data prior to multivariate modelling. *Journal of Econometrics*, *148*(1), 25–35.
- Engle, R. F., & Kozicki, S. (1993). Testing for common features. *Journal of Business & Economic Statistics*, *11*(4), 369–380.
- Francis, N., Owyang, M. T., & Savascin, O. (2017). An endogenously clustered factor approach to international business cycles. *Journal of Applied Econometrics*, *32*(7), 1261–1276.
- Fröhwrth-Schnatter, S., & Kaufmann, S. (2008). Model-based clustering of multiple time series. *Journal of Business & Economic Statistics*, *26*(1), 78–89.
- Funashima, Y. (2017). Time-varying leads and lags across frequencies using a continuous wavelet transform approach. *Economic Modelling*, *60*(1), 24–28.
- Geiger, D., Gupta, A., Costa, L. A., & Vlontzos, J. (1995). Dynamic programming for detecting, tracking, and matching deformable contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *17*(3), 294–302.
- Gullo, F., Ponti, G., Tagarelli, A., Tradigo, G., & Veltri, P. (2012). A time series approach for clustering mass spectrometry data. *Journal of Computational Science*, *3*(5), 344–355.
- Hamilton, J. D., & Owyang, M. T. (2012). The propagation of regional recessions. *Review of Economics and Statistics*, *94*(4), 935–947.
- Katris, C. (2020). Prediction of unemployment rates with time series and machine learning techniques. *Computational Economics*, *55*(2), 673–706.
- Keogh, E. J & Pazzani, M. J. (2000). Scaling up dynamic time warping for datamining applications. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 285–289).
- Keogh, E. J & Pazzani, M. J. (2001). Derivative dynamic time warping. In *Proceedings of the 2001 SIAM international conference on data mining* (pp. 1–11). SIAM.
- Liao, T. W. (2005). Clustering of time series data—A survey. *Pattern Recognition*, *38*(11), 1857–1874.
- MacQueen, J., et al. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, *1*(14), 281–297.
- Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87–106.
- Müller, M. (2007). Dynamic time warping. *Information retrieval for music and motion* (pp. 69–84). Berlin: Springer.
- Papageorgiou, T., Michaelides, P. G., & Milios, J. G. (2017). Business cycles synchronization and clustering in Europe (1960–2009). *Journal of Economics and Business*, *62*(5), 419–470.
- Petitjean, F., Ketterlin, A., & Gancarski, P. (2011). A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognition*, *44*(3), 678–693.
- Raihan, T. (2017). Predicting US recessions: A dynamic time warping exercise in economics. Unpublished working paper.

- Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1), 43–49.
- Sejdic, E., Djurovic, I., & Jiang, J. (2009). Time-frequency feature representation using energy concentration: An overview of recent advances. *Digital Signal Processing*, 19(1), 153–183.
- Sullivan, M. K., & Miller, A. (1996). Segmenting the informal venture capital market: Economic, hedonistic, and altruistic investors. *Journal of Business Research*, 36(1), 25–35.
- Wang, G.-J., Xie, C., Han, F., & Sun, B. (2012). Similarity measure and topology evolution of foreign exchange markets using dynamic time warping method: Evidence from minimal spanning tree. *Physica A: Statistical Mechanics and Its Applications*, 391(16), 4136–4146.
- Xie, Y., & Wiltgen, B. (2010). Adaptive feature based dynamic time warping. *International Journal of Computer Science and Network Security*, 10(1), 264–273.
- Zhang, Y., & Hepner, G. F. (2017). The dynamic-time-warping-based k-means++ clustering and its application in phenoregion delineation. *International Journal of Remote Sensing*, 38(6), 1720–1736.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.