Assessment of the similarity between the phases of the time series

Michał Bernardelli

Konferencja Naukowa „Metodologia Badań Statystycznych"
Warszawa, 1st-3rd July, 2023

Presentation outline

SGH

www.sgh.waw.pl

Goal and motivation

High-level goals

- To monitor the economic situation and provide some warning signals in advance.

- To learn from experience about the past events

- To find the similarities between the time series and between the phases of the time series (macroeconomic, financial, banking, etc.)

Main difficulties

- High data volatility

- Different lengths of the phases

- Hard to find the measure which captures the alternations between leading and lagging relationships

SGH

Time series similarity measures
(according to ChatGPT)

- **Euclidean Distance**: It measures the straight-line distance between two points in a multi-dimensional space and is a simple measure of similarity.

- **Correlation**: It measures the relationship between two series and is used to determine the strength and direction of the relationship.

- **Cosine Similarity**: It is a measure of similarity between two non-zero vectors of an inner product space. It measures the cosine of the angle between two vectors.

- **Jaccard Similarity**: It is a measure of similarity between two sets and is defined as the size of the intersection divided by the size of the union of the sets. It is commonly used for comparing sequences with categorical variables.

SGH

Dynamic Time Warping (DTW)

History

Introduced independently in (Vintsyuk, 1968) and (Sakoe & Chiba, 1978) as speech application.

Idea

Calculate a dissimilarity measure between two points in the time series and find the optimal alignment between them to minimise the overall distance.
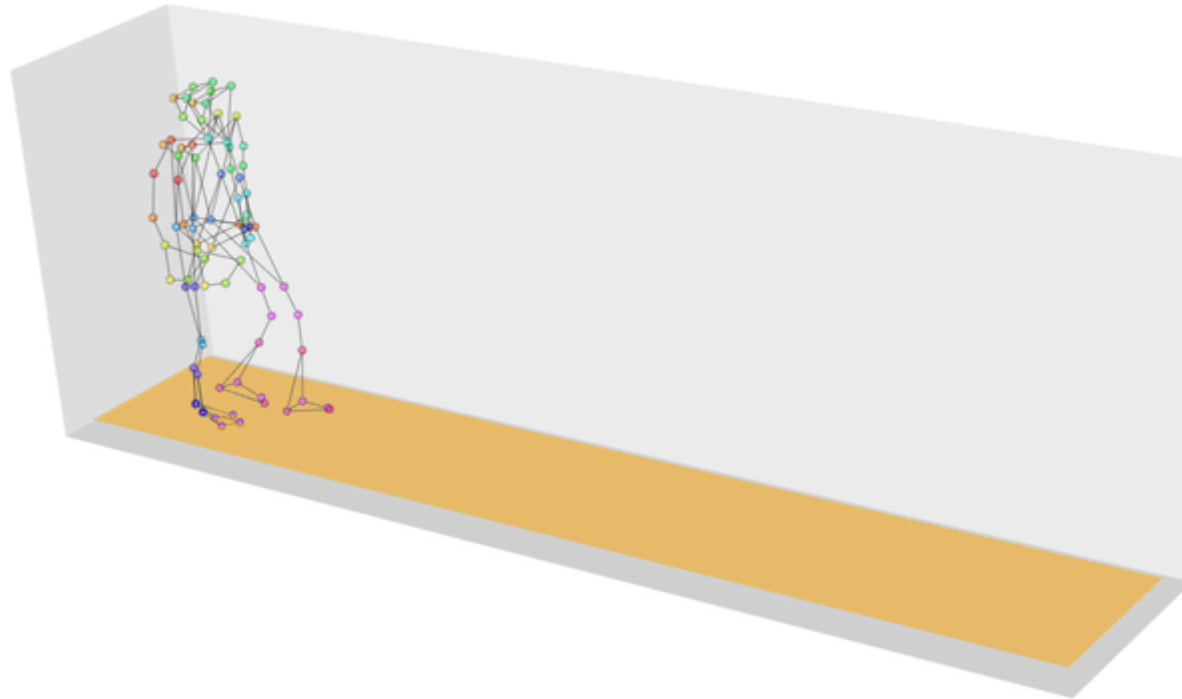
Characteristics

A powerful tool for analysing time-series data because of the ability to handle varying lengths and speeds and incorporate domain-specific.
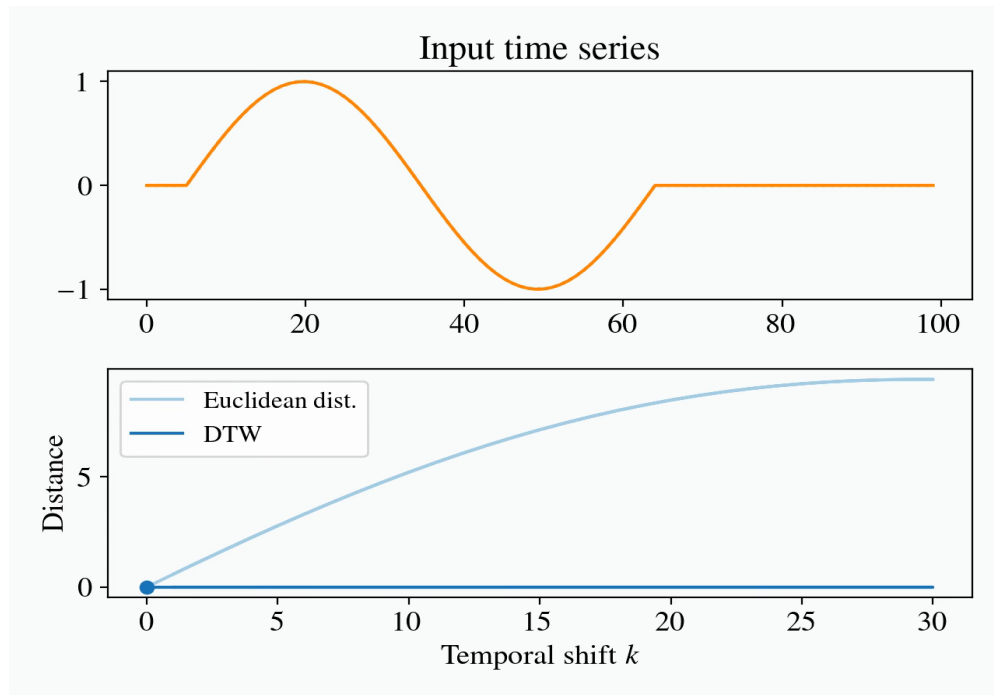
SGH

Dynamic Time Warping - applications

- **Speech recognition**: align speech recordings to a reference template to improve speech recognition systems' accuracy.

- **Music analysis**: align different performances of a piece of music and to identify similarities and differences between them.

- **Robotics**: align sensor data from robots, such as accelerometer and gyroscope readings, to recognise and track gestures.

- **Bioinformatics**: align biological sequences, such as DNA and proteins, to differentiate or compare them.

- **Monitoring**: align time-series data from various sensors, to detect anomalies in the data and monitor the performance of systems

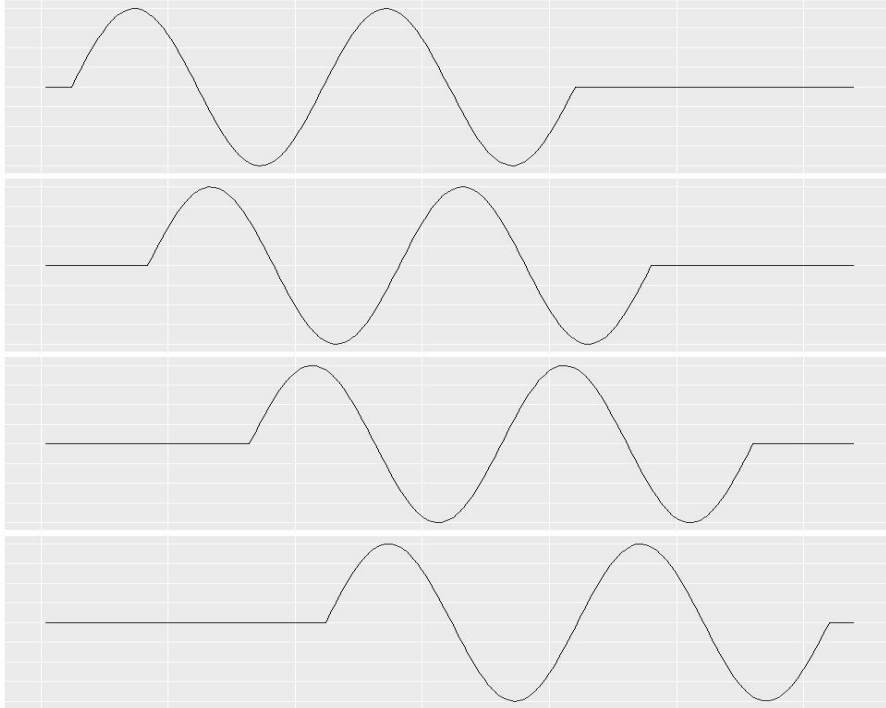**SGH**

Dynamic Time Warping – example

walking sequence recorded using a motion-capture system
source: https://en.wikipedia.org/wiki/Dynamic_time_warping

Dynamic Time Warping – example

Euclidean distance vs DTW:

DTW is invariant to time shifts between series

Source: https://rtavenar.github.io/blog/dtw.html

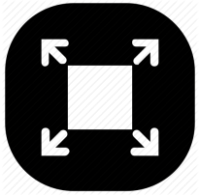**SGH**

Dynamic Time Warping – example

Time series shifted in time relative to each other:
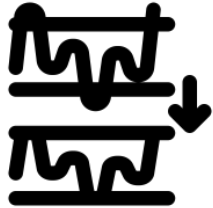
DTW measure = 0

Similarities across business cycles

Two fundamental problems
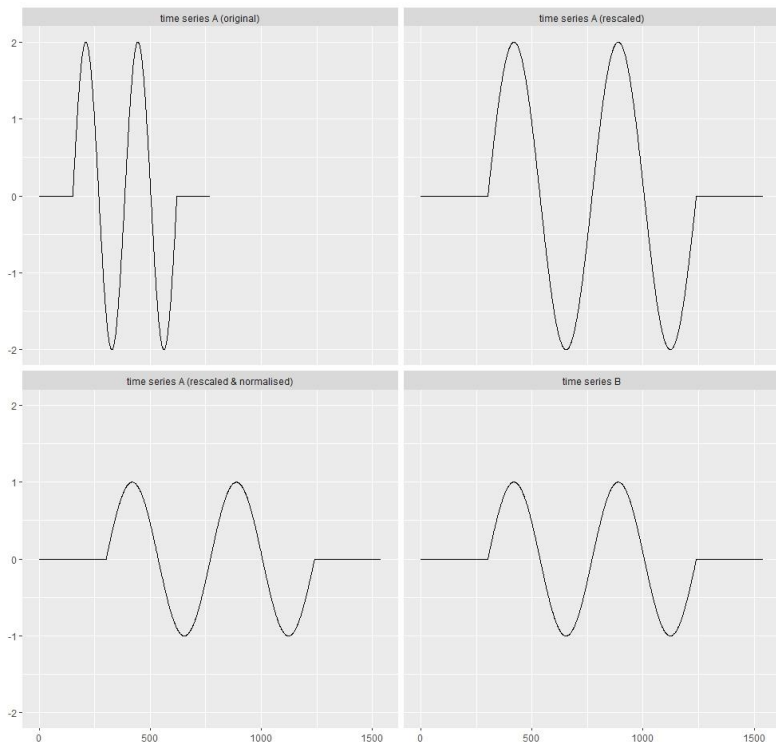
The phases of the time series have different lengths
➔ **SCALING**

The peaks and troughs have different amplitudes
➔ **NORMALISATION**

Similarities across business cycles based on DTW

time series A (original)

time series A (rescaled)

time series A (rescaled & normalised)

time series B

Time series A follows the same pattern as time series B, but twice faster and with twice the magnitude.

1. Scaling

   time series A became equal in length to time series B

2. Normalisation

$$\widetilde{x}_t = \frac{x_t}{\max_{s}|x_s|}$$

DTW measure after normalisation = 1.56

DTW measure without normalisation = 416.76

SGH

RIED ICI

Industrial Confidence Index (ICI) computed by the Research Institute for Economic Development (RIED) at SGH Warsaw School of Economics

- monthly business tendency survey in the Polish manufacturing industry

- eight questions about production volume, total and export orders, finished goods inventory, selling prices, employment, financial standing, and the general economic situation in Poland

- two versions: retrospective about the current and perspective about the future

- balance is calculated as the difference between the percentages of positive and negative answers

- RIED ICI is computed as the average of three balances: production expectations, the current volume of total orders, and finished goods inventory (with a negative sign)
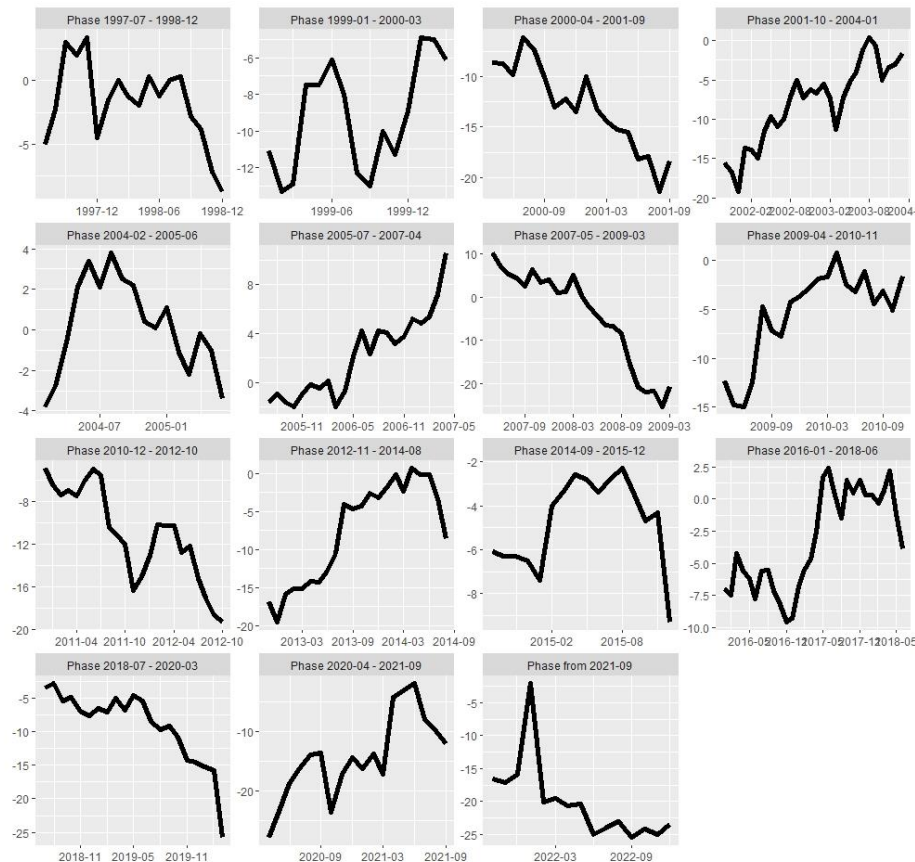
Data characteristics
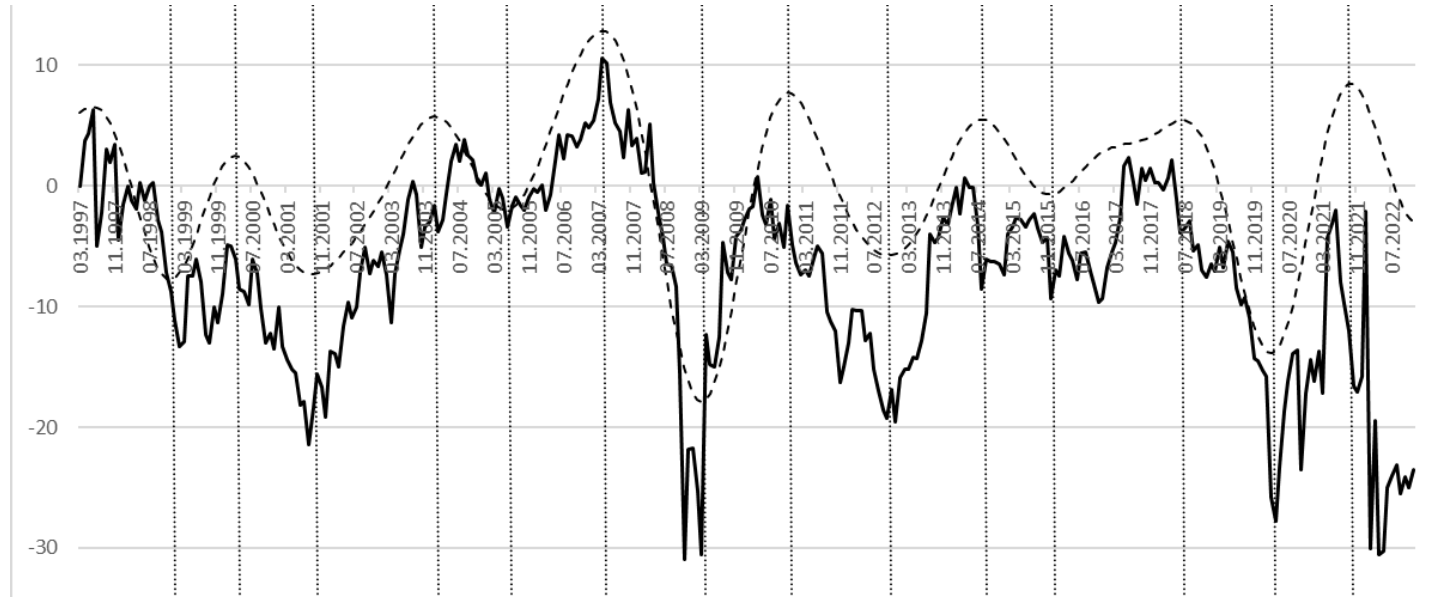

RIED ICI (solid line) along with the cyclical component (dashed line) and turning points (vertical dotted lines) from March 1997 to December 2022

Phases of business cycles identified by RIED ICI:

- different lengths (from 15 to 30 months)

- large range of quoted values (amplitudes from -27.8 to 10.6)

SGH

Phases of business cycles identified by RIED ICI after scaling and normalisation
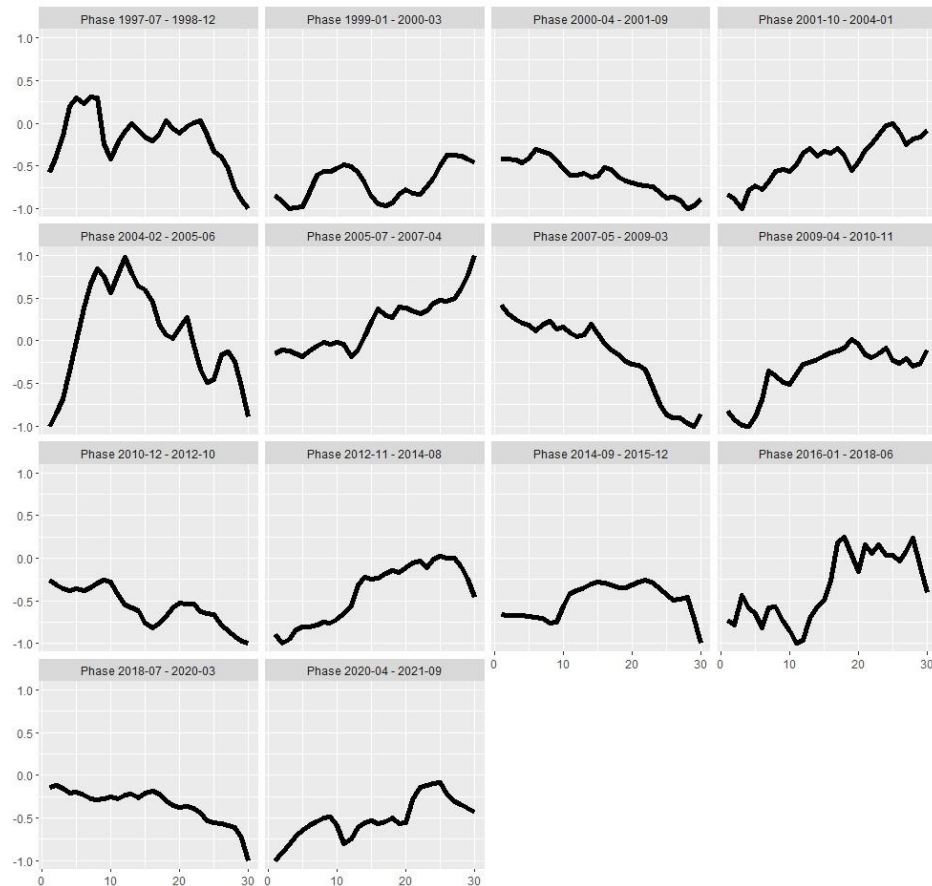
	1997-07 - 1998-12	1999-01 - 2000-03	2000-04 - 2001-09	2001-10 - 2004-01	2004-02 - 2005-06	2005-07 - 2007-04	2007-05 - 2009-03	2009-04 - 2010-11	2010-12 - 2012-10	2012-11 - 2014-08	2014-09 - 2015-12	2016-01 - 2018-06	2018-07 - 2020-03	2020-04 - 2021-09
1997-07 - 1998-12	0.0	14.3	8.0	9.1	8.6	20.9	6.4	9.3	7.9	7.9	7.8	6.7	5.7	7.2
1999-01 - 2000-03	14.3	0.0	8.8	6.2	19.6	34.7	25.7	7.7	11.1	7.3	5.4	8.8	15.4	3.5
2000-04 - 2001-09	8.0	8.8	0.0	15.3	18.0	38.8	11.8	15.9	2.3	13.9	4.9	15.6	4.2	11.4
2001-10 - 2004-01	9.1	6.2	15.3	0.0	12.5	17.7	28.4	2.7	15.9	2.3	6.5	5.5	14.1	4.0
2004-02 - 2005-06	8.6	19.6	18.0	12.5	0.0	20.0	10.5	12.1	17.5	11.6	15.8	10.3	15.4	14.1
2005-07 - 2007-04	20.9	34.7	38.8	17.7	20.0	0.0	29.0	16.0	36.9	19.5	29.7	16.5	26.1	24.4
2007-05 - 2009-03	6.4	25.7	11.8	28.4	10.5	29.0	0.0	22.6	9.1	24.4	20.3	22.2	6.4	26.9
2009-04 - 2010-11	9.3	7.7	15.9	2.7	12.1	16.0	22.6	0.0	15.3	3.0	6.7	5.7	11.2	3.6
2010-12 - 2012-10	7.9	11.1	2.3	15.9	17.5	36.9	9.1	15.3	0.0	13.9	6.5	14.7	2.7	11.6
2012-11 - 2014-08	7.9	7.3	13.9	2.3	11.6	19.5	24.4	3.0	13.9	0.0	5.9	5.4	11.5	2.5
2014-09 - 2015-12	7.8	5.4	4.9	6.5	15.8	29.7	20.3	6.7	6.5	5.9	0.0	8.1	8.2	4.4
2016-01 - 2018-06	6.7	8.8	15.6	5.5	10.3	16.5	22.2	5.7	14.7	5.4	8.1	0.0	14.1	5.4
2018-07 - 2020-03	5.7	15.4	4.2	14.1	15.4	26.1	6.4	11.2	2.7	11.5	8.2	14.1	0.0	12.4
2020-04 - 2021-09	7.2	3.5	11.4	4.0	14.1	24.4	26.9	3.6	11.6	2.5	4.4	5.4	12.4	0.0

DTW distances between pairs of business cycle phases identified by RIED ICI.

The closer the colour is to red, the smaller the distance.
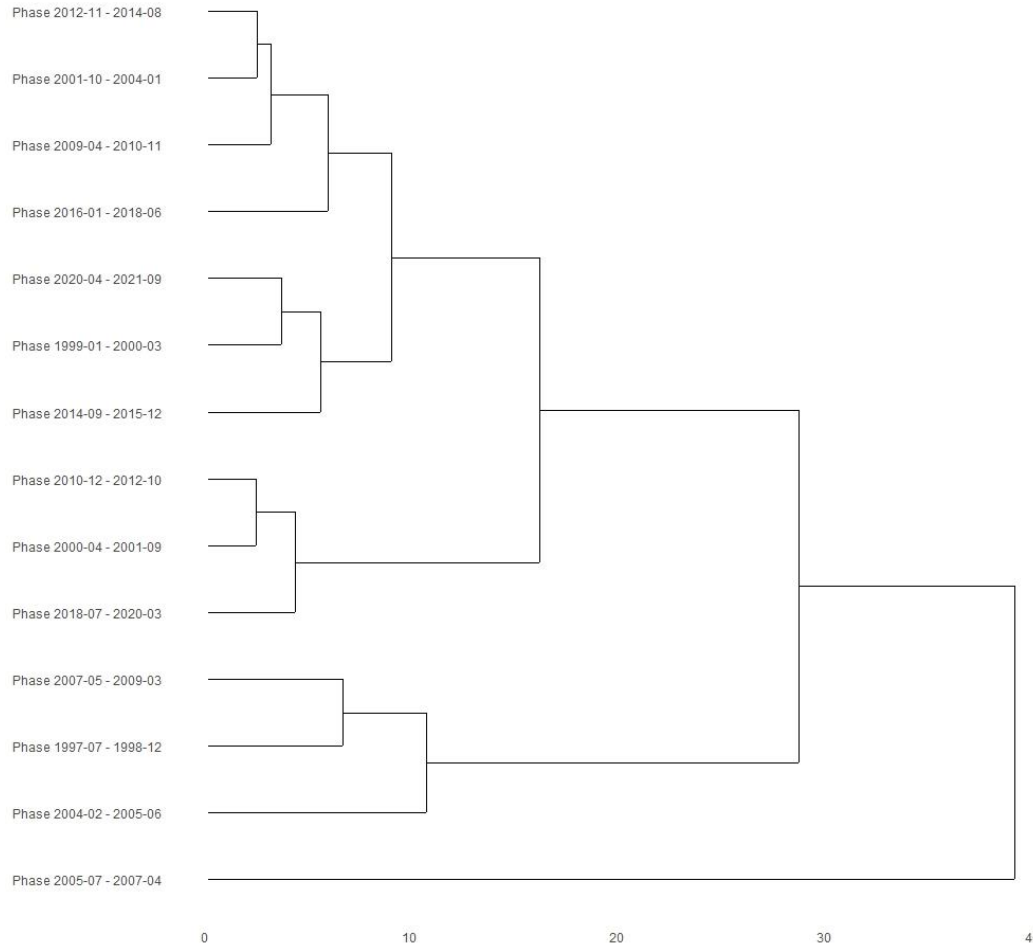
Hierarchical cluster analysis using a set of dissimilarities between business cycle phases measured by the DTW distance.

DTW distances for phases:

- 2001-10 - 2004-01

- 2009-04 - 2010-11

- 2012-11 - 2014-08

are small.

Empirical analysis limitations

- It is not an algorithm for identifying turning points but to indicate an objective method of quantifying changes in the business cycle.

- Time intervals bounded by turning points identified by RIED are not optimal concerning the DTW measure.

- The peak-through and through-peak phases can be separated in the analysis.
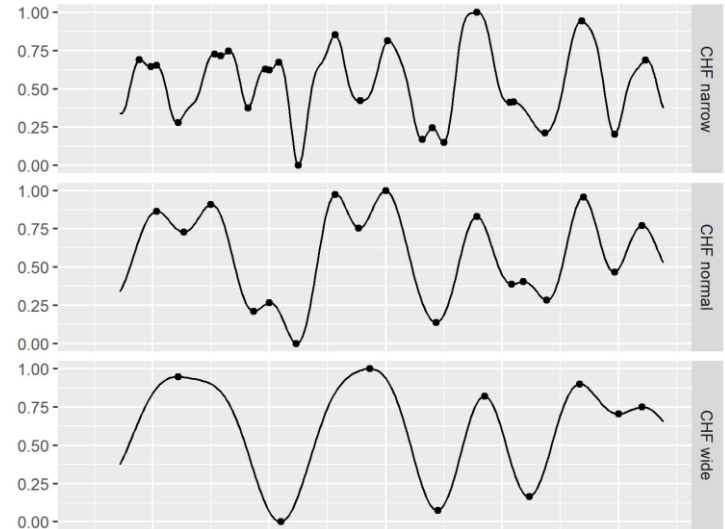
Different turning point datings depending on the parameters of the Christiano-Fitzgerald filter (used in RIED ICI).

SGH

Concluding remarks

- DTW can capture alternations between leading and lagging relationships

- The DTW concept is suitable for any length of time series

- Easy interpretability

- DTW similarity method is flexible regarding low requirements compared to the classical econometric methods

- DTW approach is universal – may be applied for analysis of various input data (macroeconomic, financial, banking, etc.)

- Allow comparing macroeconomics time series with each other and between different phases

SGH

Contact

dr hab. Michał Bernardelli, prof.

SGH

michal.bernardelli@sgh.waw.pl

Thank you for your attention!