ECONOMETRICS. EKONOMETRIA
Advances in Applied Data Analysis
Year 2019, Vol. 23, No. 2
ISSN 1507-3866; e-ISSN 2449-9994

CLUSTERING COMPANIES LISTED ON THE WARSAW STOCK EXCHANGE ACCORDING TO TIME-VARYING BETA

Piotr Szczepocki
University of Lodz, Lodz, Poland
e-mail: piotr.szczepocki@uni.lodz.pl
ORCID: 0000-0001-8377-3831

DOI: 10.15611/eada.2019.2.05
JEL Classification: C38, C58

Abstract: The beta parameter is a popular tool for the evaluation of portfolio performance. The Sharpe single-index model is a simple regression model in which the stock's returns are regressed against the returns of a broader index. The beta parameter is a measure of the strength of this relation. Extensive recent research has proved that the beta is not constant in time and should be modelled as a time-variant coefficient. One of the most popular methods of the estimation of a time-varying beta is the Kalman filter. As the output of the Kalman filter, one obtains a sequence of the estimates of a time-varying beta. This sequence shows the historical dynamics of sensitivity of a company's returns to the variations of market returns. The article proposes a method of clustering companies listed on the Warsaw Stock Exchange according to time-varying betas.

Keywords: time series clustering, cluster analysis, time-varying beta.

1. Introduction

The Sharpe single-index model [Sharpe 1964] is very popular both among financial practitioners and theoreticians. A crucial parameter in this model is the beta which shows the relation between asset's returns and market portfolio returns. For investors, betas are one of the systematic risk measures [Dębski et al. 2017]. Precise estimates for this parameter are crucial in many financial applications, including asset pricing and risk management.

The primary assumption in the single-index model is that beta is constant in time. However, many authors questioned these assumptions and showed strong empirical findings against constant betas. Consequently, static betas have been losing out in favour of time-varying betas [Andersen et al. 2006].

Time-varying betas yield extra knowledge contrary to static betas: the dynamics of betas. The aim of the paper is to cluster the major companies listed on the Warsaw

Stock Exchange according to the dynamics of betas. The clustering of economic and financial time series is a quite new, rapidly developing statistical tool which may be used to identify structural similarities and stable dependencies in economic processes for risk and investment management [Focardi, Fabozzi 2004]. Time series clustering has proved its usefulness in many areas of economics such as: analysis of personal income patterns [Bagnall et al. 2003], finding seasonality patterns in retail [Kumar, Patel 2002] and identifying patterns in macroeconomic time series [Augustyński, Laskoś-Grabowski 2018]. Finding clusters of companies which have similar dynamics of sensitivity to the variations of market returns may be advantageous in portfolio selection and risk management, because this may be used to diversify risk. There are a few works that cluster companies on the basis of time series of prices (e.g. Fu et al. [2001], Aghabozorgi and Teh [2014], Marvin [2015], Korzeniewski [2017; 2018]) but to the best of the authors' knowledge, this is the first work that attempts to cluster companies on the basis of time series of one of the risk measures – betas.

The contribution of the article to the literature is to propose the methodology of clustering time-varying betas. The article is organized as follows: Section 2 is an overview of previous related literature, Section 3 introduces the data used in the empirical example, Section 4 describes the research methodology, in Section 5 we present the results, and Section 6 concludes.

2. Literature review

It seems that no economic variable is constant over a long period of time. Thus the stability of betas has been constantly questioned. Blume [1971] studied the stationarity of betas over a very long period, from 1926 to 1968, and found that "betas tend to regress towards the means with this tendency stronger for lower risk portfolios than the higher risk portfolios". Whereas Baesel [1974] concluded that "the stability of the beta is dependent upon both the estimation interval used and upon the extremity of the beta chosen". Gonedes [1973] found that the optimal estimation interval is seven years, while for Alexander and Chervany [1980] that is generally for four to six years. Fabozzi and Francis [1977] estimated and tested the stability of betas over the bull and bear markets and found no evidence supporting beta instability. Later, the results of Kim and Zumwolt [1979] and Chen [1982] opposed this conclusion. They found that the decomposition of total systematic risk into upward and downward leads to models (time invariant parameters [Kim, Zumwolt, 1979] or time variant parameters [Chen 1982]) that have greater prognostic strength. Huang [2000] used the two-state, first-order Markov switching method introduced by Hamilton [1988] to model betas to be drawn from two different regimes, e.g. a high-risk state and low-risk state.

Since the work by Fabozzi and Francis [1978], a vast amount of literature has been devoted to different approaches to estimating time-varying betas. The most common

Table 1. List of companies included in the research

No.	Company name	Ticker	Index	No.	Company name	Ticker	Index
1	Asseco Poland SA	ACP	mWIG40	24	KGHM Polska Miedź SA	KGH	WIG20
2	Amica SA	AMC	mWIG40	25	Grupa Kęty SA	KTY	mWIG40
3	Grupa Azoty SA	ATT	mWIG40	26	LC Corp SA	LCC	mWIG40
4	Budimex SA	BDX	mWIG40	27	LPP SA	LPP	WIG20
5	Bank Handlowy SA	BHW	mWIG40	28	Grupa LOTOS SA	LTS	WIG20
6	Boryszew SA	BRS	mWIG40	29	LW Bogdanka SA	LWB	mWIG40
7	BZ WBK SA	BZW	WIG20	30	mBank SA	MBK	WIG20
8	Inter Cars SA	CAR	mWIG40	31	Bank Millennium SA	MIL	mWIG40
9	CCC SA	CCC	WIG20	32	Netia SA	NET	mWIG40
10	CD Projekt SA	CDR	WIG20	33	Orange Polska SA	OPL	WIG20
11	Ciech SA	CIE	mWIG40	34	Orbis SA	ORB	mWIG40
12	CI Games SA	CIG	mWIG40	35	Bank Pekao SA	PEO	WIG20
13	ComArch SA	CMR	mWIG40	36	Pfleiderer Group SA	PFL	mWIG40
14	Cyfrowy Polsat SA	CPS	WIG20	37	Polska Grupa Energetyczna SA	PGE	WIG20
15	AmRest Holdings SE	EAT	mWIG40	38	Polskie Górnictwo Naftowe i Gazownictwo SA	PGN	WIG20
16	Enea SA	ENA	mWIG40	39	Polski Koncern Naftowy Orlen SA	PKN	WIG20
17	Eurocash SA	EUR	WIG20	40	PKO Bank Polski SA	PKO	WIG20
18	Famur SA	FMF	mWIG40	41	Polimex-Mostostal SA	PXM	mWIG40
19	Forte SA	FTE	mWIG40	42	Powszechny Zakład Ubezpieczeń SA	PZU	WIG20
20	Getin Noble Bank SA	GNB	mWIG40	43	Sanok Rubber Company SA	SNK	mWIG40
21	Globe Trade Centre SA	GTC	mWIG40	44	Stalprodukt SA	STP	mWIG40
22	ING Bank Śląski SA	ING	mWIG40	45	Wawel SA	WWL	mWIG40
23	Kernel Holding SA	KER	mWIG40				

Source: own work.

approaches are: betas estimated in the rolling window within a linear regression, multivariate GARCH models (MGARCH), the Kalman filter and realized betas derived from realized covariance and variance [Andersen et al. 2006]. However, there are still many other propositions, e.g. Chen and Lee [1982] introduced Bayesian inference, Ferreira et al. [2011] proposed a two-stage nonparametric approach, Cai et al. [2015] proposed a functional coefficient regression technique. All these methods of estimation have some advantages and drawbacks. A few authors compared the competitive approaches on different markets [Brooks, Faff, McKenzie 1998; Lie, Brooks, Faff 2000] and found that the Kalman filter performed at least equally or even better than the MGARCH specifications. As far as the Polish capital market is concerned, Będowska-Sójka [2017] compared the beta coefficients obtained from MGARCH and the Kalman filter on data of weekly frequency for stocks quoted on the Warsaw Stock Exchange in terms of in-sample predictive accuracy and did not find statistical difference between the accuracy of the DCC MGARCH model and the Kalman filter. In this paper we use the Kalman filter approach to the estimation of time varying betas.

3. Data

For the purposes of this research we took into consideration the 60 largest companies listed on the Warsaw Stock Exchange[1] from two capitalization-weighted stock market index: the WIG20 (the 20 biggest and the most liquid companies of the WSE Main List) and the mWIG40 (the mid-cap index that consists of 40 medium-size companies of the WSE Main List). We decided to use weekly stock returns from the period 2010-06-04 to 2017-12-29. As market returns we chose the Warsaw Stock Exchange Index WIG (*Warszawski Indeks Giełdowy*). Price data was obtained from the Stooq database (https://stooq.pl/). Only 44 companies out of the 60 which were constantly quoted for the whole period were included in the research. Table 1 shows the list of examined companies. The whole sample consists of 396 weekly returns. We transformed the returns into percentage logarithmic returns for further work.

4. Research methodology

We start with the beta estimation technique. The Kalman filter requires rewriting the dynamic system in the state-space representation, which consists of two equations: measurement and transition. In the case of time, the time-varying beta estimation former equation is the security characteristic line with the time-varying beta coefficient:

$$R_{it} = \alpha_i + \beta_{it} R_{Mt} + \varepsilon_t, \tag{1}$$

[1] As of 20 July 2018.

where R_{it} and R_{Mt} are the return on asset i and on the market portfolio at time t, respectively. The error term ε_t is a zero-mean normally and independently distributed with constant variance σ_ε^2. The transition equation determines how beta changes over time. In the literature there seems to be no agreement as to what form the transition equation should take. There are many proposals e.g. AR(1), mean-revision, random coefficient, however the most popular is the assumption that a time-varying beta follows random walk:

$$\beta_{it} = \beta_{it-1} + \eta_t, \tag{2}$$

where η_t is a zero-mean normally and independently distributed with constant variance σ_η^2 (ε_t and η_t are independent variables). There is a vast amount of literature that supports this form of transition equation, inclusing: Faff et al. [2000], Ebner and Neumann [2005], Choudhry and Wu [2008], Das and Ghoshal [2010], Kurach and Stelmach [2014], Będowska-Sójka [2017]. Random walk assumption means that any shocks to beta persist forever.

As the output of the Kalman filter, one obtains a sequence $\left(\hat{\beta}_{it}\right)_{t=1}^{T}$ containing the filtered state variables:

$$\hat{\beta}_{it} = E\left[\beta_{it} \mid \left(R_{it}, R_{Mt}\right)_{1:t}\right], \tag{3}$$

for t = 1,..,T (T is length of both time series R_{it} and R_{Mt}). This sequence $\left(\hat{\beta}_{it}\right)_{t=1}^{T}$ shows dynamics of sensitivity of returns on asset I to changes in the market portfolio returns and it is the base of time series clustering.

We use standardisation estimates of time-varying betas before clustering:

$$\tilde{\beta}_{it} = \frac{\hat{\beta}_{it} - \overline{\beta}_i}{s_{\beta_i}}. \tag{4}$$

According to many recent works concerned with time series clustering, standardisation is an essential pre-processing step which allows to focus on the structural similarities rather than the similarities that come from amplitude [Paparrizos, Gravano 2015].

Time series clustering is a type of clustering algorithm which handles dynamic data. The most important element in time series clustering is to choose the dissimilarity or distance measure between two time series. Aghabozorgi et al. [2015] distinguish three types of distance measures: measures based, feature-based and model-based. For the scope of this paper, we focus on the first. In the shape based approach, shapes of two time-series are compared to find similarity of patterns in the presence of a variety of distortions, e.g. differences in amplitude and phase. As the notion of shape is not precisely defined, numerous distance measures have been proposed (an overview of literature on time-series distance measures can be found in [Montero and Vilar 2014] and [Aghabozorgi et al. 2015]). In this paper we propose to use Dynamic Time Warping (DTW) as the dissimilarity measure. In the context of shape-based time-series clustering it is very common to use DTW [Aghabozorgi et al. 2015].

The popularity of this measure results from its resistance to transformations such as shifting and/or scaling.

DTW is an algorithm for measuring dissimilarity between two time-series that tries to find an optimal match between them under certain constraints. This optimal match minimises cost, where the cost is computed on the basis of the sum of differences (absolute or the Euclidean distance) between the considered time series. We will briefly summarize the algorithm below. The paper by Giorgino [2009] gives a detailed description of the DTW.

We assume that we want to compare two time series: $\boldsymbol{X} = (x_1, x_2, \ldots, x_T)$ and $\boldsymbol{Y} = (y_1, y_2, \ldots, y_T)$[2]. In the first step in DTW a Local Cost Matrix (LCM) is created. For each element (i, j) of the matrix a norm (absolute, Euclidean distances or the more general l_p norm) between x_i and y_j are computed. Then, the DTW algorithm finds the path through the cost matrix, starting at LCM(1, 1) and finishing at LCM(T,T) (boundary conditions), aggregating the cost at each step. Apart from boundary conditions, the path needs to meet more conditions such as monotonicity (both the *i* and *j* indexes may either stay the same or increase, but they may not decrease) or continuity (both indexes can only increase by 1 on each step).

Defining by $\varphi = \{(1,1), \ldots, (T,T)\}$ the set of all the points that belong to a path, the final distance may be computed by the equation

$$d_\varphi(X,Y) = \sum_{k=1}^{T} \frac{d(\varphi_x(k), \varphi_y(k)) m_\varphi(k)}{M_\varphi}, \tag{5}$$

where d is l_p norm, $m_\phi(k)$ is a per-step weighting coefficient and M_ϕ is the corresponding normalization constant. Both $m_\phi(k)$ and M_ϕ depend on the chosen step patterns. Step patterns[3] are various modifications of plain DTW that control the possible routes of the warping paths, especially to avoid duplication of elements (e.g. a single time point in X match multiple (consecutive) elements in *Y*). Finally, the DTW similarity measure between the two time-series $\boldsymbol{X}$ and $\boldsymbol{Y}$ minimises $d_\phi(X, Y)$ for all possible paths

$$DTW(X,Y) = \min_\varphi d_\varphi(X,Y). \tag{6}$$

In the research we use the dtw_basic function from **R** package **dtwclust** [Giorgino 2009] to calculate the dissimilarity measure between the time-series of time varying betas with default settings: l_1 norm (absolute distance), symmetric2 step pattern (allows an unlimited number of elements of one of the time-series to be matched to a single element of the other time-series) without normalisation or global constraints[4].

---

[2] The DTW distance may be potentially used with time-series of different lengths.

[3] More on step patterns can be found in Sakoe and Chiba [1978].

[4] Global constraints forbid paths to enter some region of the Local Cost Matrix, especially to move too far from the diagonal (Sakoe-Chiba band [Sakoe and Chiba 1978]).

On the basis of the DTW dissimilarity measure, we create a distance matrix that represents the distance (dissimilarity) between each pair of time-series of time-varying betas. We use the agglomerative hierarchical clustering method with Ward linkage to build a hierarchy of clusters and present the result of clustering as a dendrogram. To determine the number of clusters we use the Caliński-Harabasz index [1974] which is one of the best indices assessing the quality of classification [Walesiak 2009; Bryja 2012; Korzeniewski 2014]. This index is defined by the formula:

$$CH(k) = \frac{\frac{1}{k} tr\left(B_k\right)}{\frac{1}{n-k} tr\left(W_k\right)}, \tag{7}$$

where B_k – between group variance matrix, W_k – within group variance matrix, n – number of data objects (see Gatnar and Walesiak [2004] for more details). The optimal number of clusters is that which maximises the value of the Caliński-Harabasz index. We use the **TSclust** [Montero, Vilar 2014] **R** package to calculate the dissimilarity matrix, hclust() function from the **stats R** package to perform hierarchical clustering.

5. Empirical example

In this section we present the research results on the basis of the data presented in Section 2 and the methodology from Section 2. Figure A.1 (in the Appendix) shows standardised estimates of time-varying betas of examined companies. Figure 1 shows the resulting dendrogram. The optimal number of clusters is 12. Figure 2 and Table A.1 (in the Appendix) shows the clusters' members. Some clusters have a clear interpretation. Cluster C_8={PZU, ENA} is a cluster with companies that time-varying betas were continuously increasing (companies' returns became more sensitive to variability of market returns). On the other hand, cluster C_{12}={PXM, PEO} consists of companies that time-varying betas were decreasing (companies' returns became less prone to variability of market returns). Cluster C_{10}={CIE, BDX} indicates a pair of companies for which time varying betas were similar in spite of significant volatility. Cluster C_{12}={PXM, PEO} shows that DTW measure is resistant to transformations such as shifting: PXM has a similar trajectory of time-varying beta to PEO, but shifted by around half a year. Other clusters with more members are more difficult to interpret but still may give some insight into companies' returns historically sensitive to variability of market returns. For example, on the basis of cluster C_4={LCC, ING, FTE, ATT} one can easily check that these four companies had the peak of their sensitivity around the beginning of 2014. Companies from cluster C_6={SNK, PKO, PFL} had a gap of time-varying betas around the beginning of 2013. Quite surprisingly the clusters are not connected with sectors of the economy or companies' size (WIG20 versus mWIG40).

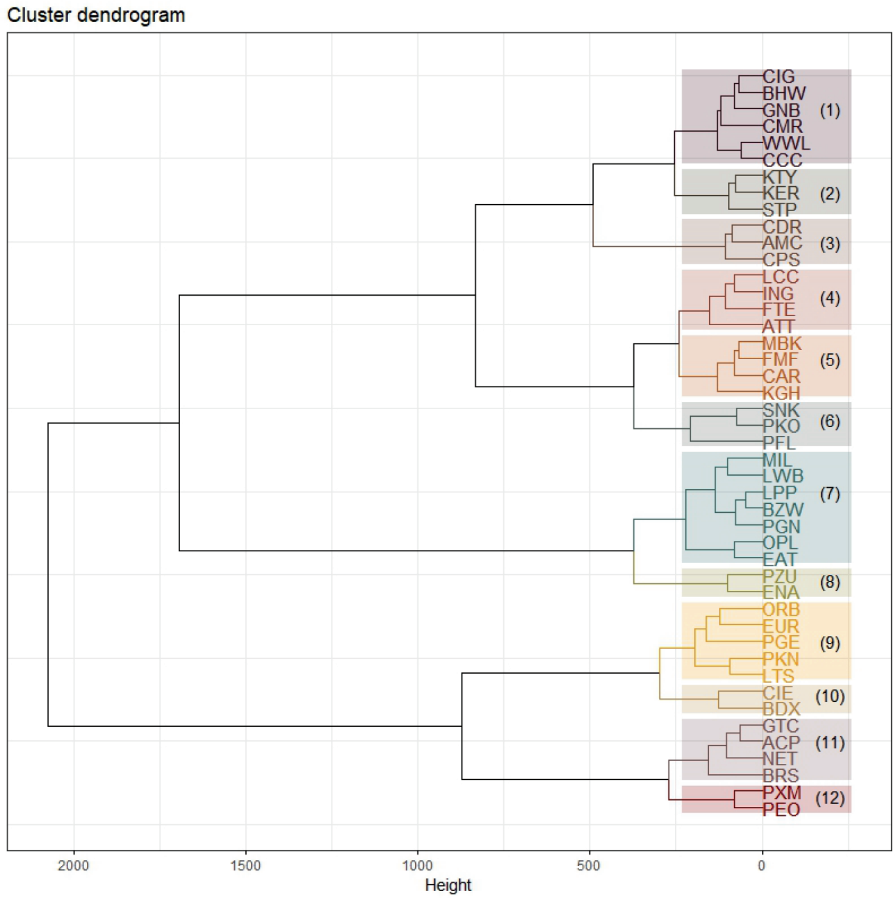

Fig. 1. Dendrogram of time-varying betas using the Ward linkage

Source: own calculations.

We extend the empirical part with a simple portfolio example. There are numerous methods to construct a stock portfolio. The classical Mean-Variance (MV) portfolio selection model of Markowitz [Markowitz 1952; 1959] is the best known. Investors' portfolios specifically reflect their own unique goals, objectives and risk tolerances.On the basis of the time-varying beta alone it is difficult to create such a portfolio because the risk is only one of a few factors. However, the presented method of clustering companies on the basis of time varying-betas may be incorporated in the process of portfolio construction to diversify exposure to market risk.

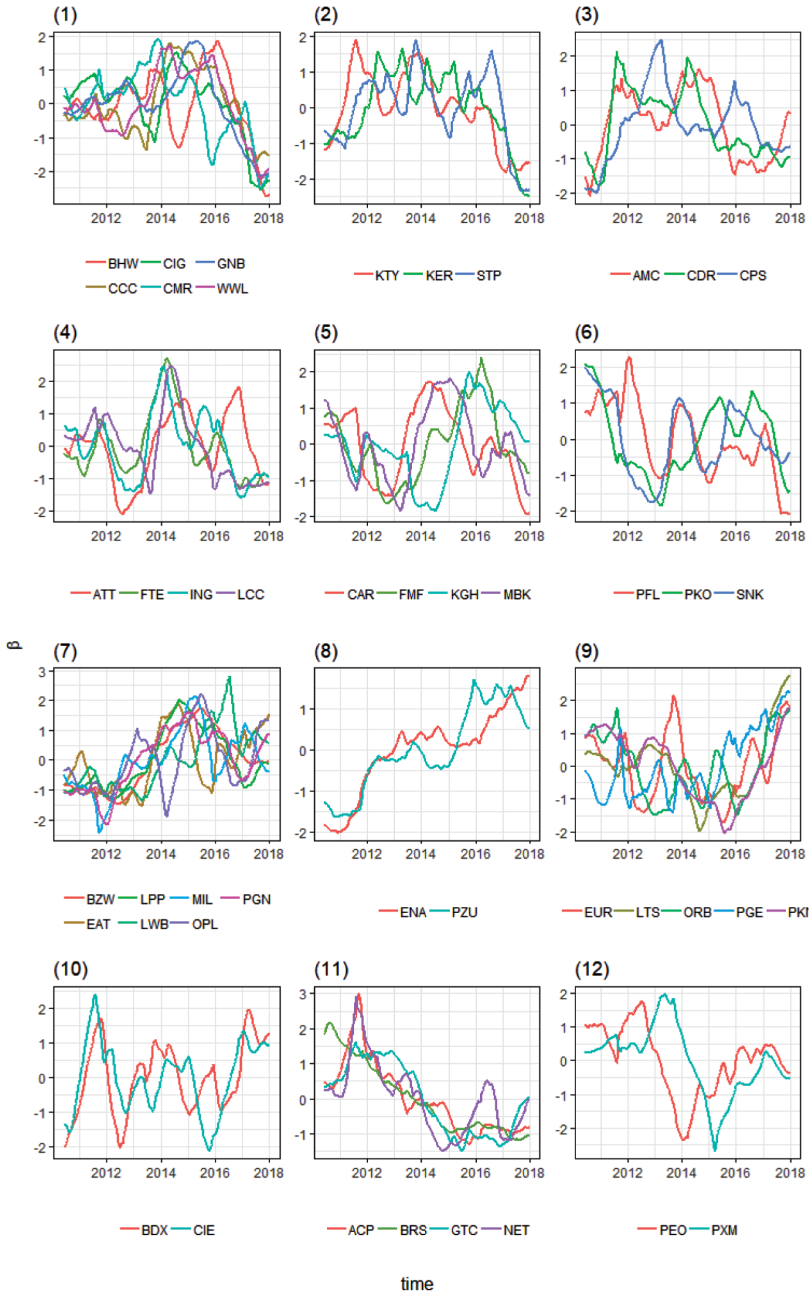

Fig. 2. Clusters of time-varying betas

Source: own calculations.

Let us consider a portfolio of twelve stocks. We take one stock from each cluster. Each stock is equally weighted in the portfolio. To reflect that investors often want to maximise the reward given the risk (as in the Mean-Variance portfolio), we pick stock with the highest mean of Treynor ratios [Treynor 1965] in the cluster. A similar strategy was used by Marvin [2015] who picked the stock from each cluster with the highest Sharpe ratio [Sharpe 1966]. The Treynor Ratio is a Return/Risk indicator given by the return earned in excess of that which could have been earned on an investment that has no risk, divided by the beta during the same period. Because we consider time-varying beta, the Treynor ratio becomes also time-varying and we take the mean value of the Treynor ratios based on the whole sample period. Following Rubaszek [2012], as the risk-free investment we consider $R_{ft} = \ln(1 + \tilde{R}_{ft} / 12)$, where $\tilde{R}_{f_t}$ is one-month spot WIBOR (Warsaw Interbank Offered Rate) at period t. Table 2 presents the portfolio of the selected company stocks and the corresponding means of the Treynor ratios.

Table 2. List of companies included in the portfolio and the corresponding means of the Treynor ratios

No. Cluster	Ticker	Mean of Treynor ratio
1	WWL	0.4598
2	KTY	0.7239
3	CDR	2.8776
4	ATT	1.0666
5	FMF	0.8812
6	SNK	0.6143
7	EAT	0.7687
8	PZU	0.3116
9	PKN	0.719
10	BDX	0.4018
11	BRS	3.3018
12	PEO	0.0336

Source: own calculations.

In Table 3 the results of the portfolio with respect to market returns (the Warsaw Stock Exchange Index WIG) are presented. The portfolio has a higher mean and cumulative sum than market returns but is more volatile. Furthermore, the portfolio has more positive returns (gains), less negative (losses) and a higher average gain (mean of positive returns). On the other hand, market returns have a smaller average loss. Figure 3 presents the curves of the cumulative sum of the portfolio and market returns.

Figure 4 presents non-standardised time-varying betas for each company's stocks and time-varying beta for the portfolio. A well-balanced portfolio in respect to the exposure to market risk should have a beta close to 1, because a beta of 1 represents the volatility of the given index used to represent the overall market, against which stocks and their betas are measured. Blume [1971] considered 1 as the "grand mean of all betas". The mean value of a time-varying beta for the portfolio is 0.901 with standard deviation 0.094. This may be considered insufficiently close to 1, but one should bear in mind that the mean value of time-varying betas of all 45 stocks that was used in research is 0.861 with standard deviation 0.114. Consequently, the portfolio of 12 stocks is better balanced in respect to exposure to market risk than the portfolio of 45 equally weighted stocks.

Table 3. Descriptive statistics of portfolio and market returns (Warsaw Stock Exchange Index WIG)

	Cumulative sum	Mean	Standard devation	Count of negative returns	Count of positive returns	Mean of negative returns	Mean of positive returns
WIG	43.141	0.109	2.061	180	216	–1.569	1.507
Portfolio	152.297	0.385	2.223	153	243	–1.701	1.698

Source: own calculations.

Fig. 3. Cumulative sum of portfolio and market returns (the Warsaw Stock Exchange Index WIG)

Source: own calculations.

This is a simple example of a portfolio of twenty equally weighted stock revealing that the presented method of clustering companies on the basis of time varying-betas may be successfully incorporated in the process of portfolio construction. However, this is only a single time-horizon example (as in the Markowitz Mean-Variance portfolio): the portfolio is constructed and evaluated at the same time horizon.

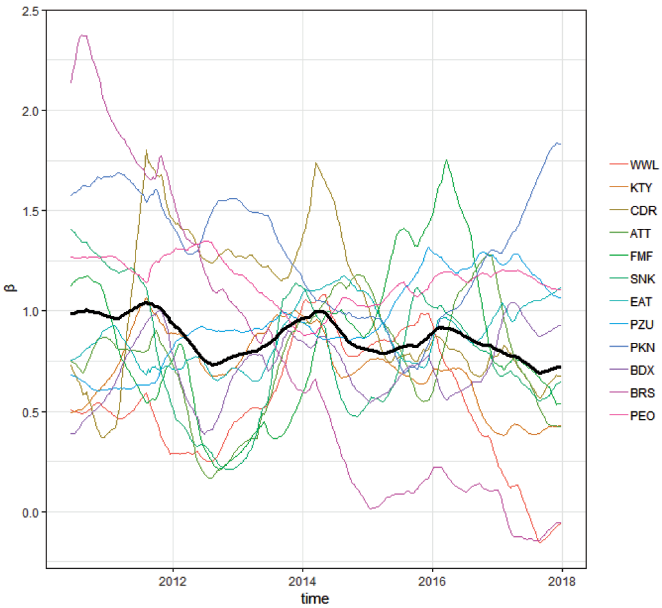

Fig. 4. Non-standardised time-varying betas for each company's stocks (coloured lines) and time-varying beta for portfolio (black line)

Source: own calculations.

6. Concluding remarks

The aim of this paper was to propose a method of clustering of the time series on the basis of time-varying betas. There is a large amount of literature on the estimation of time-varying betas, but there is still little on its practical implementation. The article tries to fill this gap. We also present an empirical example based on the Warsaw Stock Exchange and suggest some interpretation of the resulting clusters. The example is based on weekly data on moderate-length time series. However, the authors would like to emphasise two important points. Firstly, the proposed method is prone both to the length and frequency of the examined time series. This is typical problem that arises when dealing with financial data (e.g. estimation of high frequency data). The length and frequency of the time-series should be carefully chosen with regard to the considered investments. However, the Kalman filter estimates may be easily lengthened due to their recursive algorithm. Secondly, we do not believe that on the basis alone of the presented methodology one could create a reasonable portfolio. We expect only that the clustering series of time-varying betas may provide extra knowledge about the historical sensitivity to changes of market returns and may help investors to diversify their portfolios. The presented example of a portfolio of twelve equally weighted stocks proved to have a better balanced time-varying beta than a portfolio of 45 equally weighted stocks.

The main shortcoming of the study is that it only has a single time horizon. Further work should concentrate on the out-of-sample stability of the clusters and potentially on the online algorithms for the updates of clusters.

Bibliography

Aghabozorgi S., Teh Y.W., 2014, *Stock market co-movement assessment using a three-phase clustering method*, Expert Systems with Applications, vol. 41, no. 4, pp. 1301-1314.

Aghabozorgi S., Shirkhorshidi A.S., Wah T.Y., 2015, *Time-series clustering – a decade review*, Information Systems, vol. 53, pp. 16-38.

Alexander G.J., Chervany N.L., 1980, *On the estimation and stability of beta*, Journal of Financial and Quantitative Analysis, vol. 15, no. 1, pp. 123-137.

Andersen T.G., Bollerslev T., Diebold F.X., Wu G., 2006, *Realized beta: persistence and predictability*, Econometric Analysis of Financial and Economic Time Series, Emerald Group Publishing Limited, pp. 1-39.

Augustyński I., Laskoś-Grabowski P., 2018, *Clustering macroeconomic time series*, Econometrics. Advances in Applied Data Analysis, vol. 22, no. 2, 74-88.

Baesel J.B., 1974, *On the assessment of risk: some further considerations*, The Journal of Finance, vol. 29, pp. 1491-1494.

Bagnall A.J., Janacek G., De la Iglesia B., Zhang M., 2003, *Clustering Time Series from a Mixture of Polynomial Models with Discretised Data*, [in:] *Proceedings of the Second Australasian Data Mining Workshop*, pp. 105-120.

Będowska-Sójka B., 2017, *Evaluating the accuracy of time-varying Beta. The evidence from Poland*, Dynamic Econometric Models, vol. 17, pp. 161-176.

Blume M., 1971, *The assessment of risk,* The Journal of Finance, vol. 26, no. 1.

Brooks R.D., Faff R.W., McKenzie M.D., 1998, *Time-varying beta risk of Australian industry portfolios: a comparison of modeling techniques*, Australian Journal of Management, vol. 23, no. 1, pp. 1-22.

Bryja A., 2012, *Optymalizacja liczby skupień na podstawie wybranych wskaźników jakości grupowania*, Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie, vol. 892, pp. 53-67.

Caliński T., Harabasz J., 1974, *A dendrite method for cluster analysis*, Communications in Statistics-Theory and Methods, vol. 3, no. 1, pp. 1-27.

Cai Z., Ren Y., Yang B., 2015, *A semiparametric conditional capital asset pricing model*, Journal of Banking & Finance, 61, pp. 117-126.

Chen S.N., 1982, *An examination of risk-return relationship in bull and bear markets using time-varying security betas*, Journal of Financial and Quantitative Analysis, vol. 17, pp. 265-286.

Chen S.N., Lee C.F., 1982, *Bayesian and mixed estimators of time varying betas*, Journal of Economics and Business, vol. 34, no. 4, pp. 291-301.

Choudhry T., Wu H., 2009, *Forecasting the weekly time-varying beta of UK firms: comparison between the GARCH models vs the Kalman filter method*, The European Journal of Finance, vol. 15, issue 4, pp. 437-444.

Das A., Ghoshal T.K., 2010, *Market risk beta estimation using the adaptive Kalman filter*, International Journal of Engineering Science and Technology, vol. 2, no. 6, pp. 1923-1934.

Dębski W., Feder-Sempach E.M., Wójcik S., 2017, *Stabilność parametru beta dla największych spółek z rynku polskiego, niemieckiego i francuskiego – analiza porównawcza*, Annales Universitatis Mariae Curie-Skłodowska, sectio H–Oeconomia, vol. 51, no. 5.

Ebner M., Neumann T., 2005, *Time-varying betas of German stock returns*, Financial Markets and Portfolio Management, vol. 19, no. 1, pp. 29-46.

Fabozzi F., Francis J.C., 1977, *Stability tests for alphas and betas over bull and bear market conditions*, Journal of Finance, vol. 32, pp. 1093-1099.

Fabozzi F.J., Francis J.C., 1978, *Betas as a random coefficient*, Journal of Financial and Quantitative Analysis, vol. 13, pp. 101-115.

Faff R.W., Hillier D., Hillier J., 2000, *Time varying beta risk: an analysis of alternative modelling techniques*, Journal of Business Finance & Accounting, vol. 27, no. 5-6, pp. 523-554.

Ferreira E., Gil-Bazo J., Orbe S., 2011, *Conditional beta pricing models: A nonparametric approach*, Journal of Banking & Finance, vol. 35, no. 12, pp. 3362-3382.

Focardi S.M., Fabozzi F.J., 2004, *Clustering economic and financial time series: exploring the existence of stable correlation conditions*, Finance Letters, vol. 2, no. 3, pp. 1-9. Retrieved from https://www.researchgate.net/publication/2547390.

Fu T.C., Chung F.L., Ng V., Luk R., 2001, *Pattern discovery from stock time series using self-organizing maps*, [in:] *Workshop Notes of KDD 2001 Workshop on Temporal Data Mining*, pp. 26-29.

Gatnar E., Walesiak M., 2004, *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, Wydawnictwo AE we Wrocławiu, Wrocław.

Giorgino T., 2009, *Computing and visualizing dynamic time warping alignments in R: the DTW package*, Journal of Statistical Software, vol. 31, no. 7, pp. 1-24.

Gonedes N.J., 1973, *Evidence on the information content of accounting numbers: accounting-based and market-based estimates of systematic risk*, Journal of Financial and Quantitative Analysis, vol. 8, no. 3, pp. 407-443.

Hamilton J.D., 1988, *Rational-expectations econometric analysis of changes in regime: an investigation of the term structure of interest rates*, Journal of Economic Dynamics and Control, vol. 12, pp. 385-423.

Huang H-C., 2000, *Tests of regimes-switching CAPM*, Applied Financial Economics, vol. 10, pp. 573-578.

Kim K.M., Zumwalt, J.K., 1979, *An analysis of risk in bull and bear markets*, Journal of Financial and Quantitative Analysis, vol. 14, no. 5, pp. 1015-25.

Korzeniewski J., 2014, *Indeks wyboru liczby skupień w zbiorze danych*, Przegląd Statystyczny, vol. 61, no. 2, pp. 169-180.

Korzeniewski J., 2017, *Zastosowanie analizy skupień do konstruowania portfeli akcji na WGPW*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 468 (Taksonomia 28), pp. 108-115.

Korzeniewski J., 2018, *Efficient stock portfolio construction by means of clustering*, Acta Universitatis Lodziensis. Folia Oeconomica, vol. 1, pp. 85-92.

Kumar M., Patel N.R., 2002, *Clustering Seasonality Patterns in the Presence of Errors*, [in:] *Proceedings of the Eighth,* ACMSIGKDD, pp. 557-563.

Kurach R., Stelmach J., 2014, *Time-varying behaviour of sector beta risk – the case of Poland*, Romanian Journal of Economic Forecasting, vol. 17, no. 1, pp. 139-159.

Lie F., Brooks R., Faff R., 2000, *Modelling the equity beta risk of Australian financial sector companies*, Australian Economic Papers, vol. 39, no. 3, pp. 301-311.

Markowitz H. M., 1952, *Portfolio selection*, The Journal of Finance, vol. 7, no. 1, pp. 77-91.

Markovitz, H. M., 1959, *Portfolio selection: the efficient diversification of investments*, John Wiley, NY.

Marvin K., 2015, *Creating Diversified Portfolios Using Cluster Analysis*, Princeton University.

Montero P., Vilar J.A., 2014, *Tsclust: an r package for time series clustering*, Journal of Statistical Software, vol. 62, no. 1, pp. 1-43.

Paparrizos J., Gravano L., 2015, *k-shape: the efficient and accurate clustering of time series*, In Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, pp. 1855--1870.

Rubaszek M., 2012, *Modelowanie polskiej gospodarki z pakietem R*, Oficyna Wydawnicza – Szkoła Główna Handlowa.

Sakoe H., Chiba S., 1978, *Dynamic programming algorithm optimization for spoken word recognition*, IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 26, no. 1, pp. 43-49.

Sharpe W., 1964, *Capital asset prices: a theory of market equilibrium under conditions of risk*, Journal of Finance, 19, 3, pp. 425-442.

Sharpe W.F., 1966, *Mutual fund performance*, The Journal of Business, vol. 39, no. 1, pp. 119-138.

Treynor J.L., 1965, *How to rate management of investment funds*, Harvard Business Review, vol. 43, no. 1, pp. 63-75.

Walesiak M., 2009, *Analiza skupień,* [in:] *Statystyczna analiza danych z wykorzystaniem programu R*, ed. M. Walesiak, E. Gatnar, Wydawnictwo Naukowe PWN, Warszawa.

GRUPOWANIE SPÓŁEK NOTOWANYCH NA GIEŁDZIE PAPIERÓW WARTOŚCIOWYCH W WARSZAWIE WEDŁUG BET ZMIENNYCH W CZASIE

Streszczenie: Jednym z podstawowych narzędzi konstrukcji portfela akcji jest jednowskaźnikowy model Sharpe'a. Jest to model opisujący zależność pomiędzy stopami zwrotu z akcji danej spółki a czynnikiem rynkowym utożsamianym zazwyczaj z indeksem giełdowym. Miarą siły tej zależności jest parametr regresji w liniowym modelu regresji Sharpa, nazywany parametrem beta. Wiele badań wskazuje jednak, że parametr beta nie jest stabilny w czasie i do jego wyznaczenia należy użyć modeli, które umożliwiają opisanie dynamiki parametru beta w czasie. Jedną z najczęściej używanych metod do oszacowania parametru beta zmiennego w czasie jest filtr Kalmana. Jako wynik filtru Kalmana otrzymujemy szereg czasowy będący oszacowaniem parametru beta zmiennego w czasie. W artykule zostaną zaprezentowane przykłady grupowania spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie ze względu na otrzymane oszacowania parametru beta zmiennego w czasie z wykorzystaniem miary DTW (*Dynamic Time Warping*).

Słowa kluczowe: grupowanie szeregów czasowych, analiza skupień, bety zmienne w czasie, CAPM.

Appendix

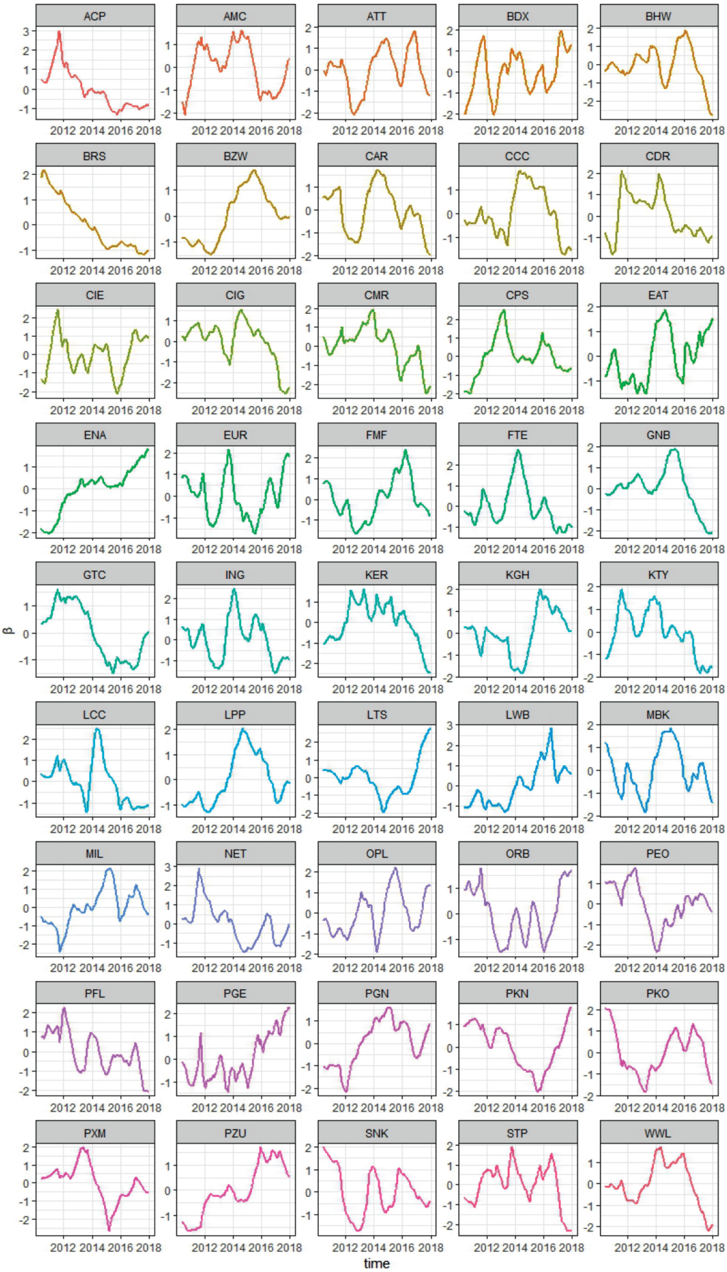

Fig. A.1. Time–varying betas filtered by the Kalman filter after standardisation

Source: own calculations.

Table A.1. List of members of each cluster

No. Cluster	Company name	Ticker	Index	No. Cluster	Company name	Ticker	Index
1	Bank Handlowy SA	BHW	mWIG40	7	BZ WBK SA	BZW	WIG20
	CCC SA	CCC	WIG20		AmRest Holdings SE	EAT	mWIG40
	CI Games SA	CIG	mWIG40		LPP SA	LPP	WIG20
	ComArch SA	CMR	mWIG40		LW Bogdanka SA	LWB	mWIG40
	Getin Noble Bank SA	GNB	mWIG40		Bank Millennium SA	MIL	mWIG40
	Wawel SA	WWL	mWIG40		Orange Polska SA	OPL	WIG20
2	Kernel Holding SA	KER	mWIG40		Polskie Górnictwo Naftowe i Gazownictwo SA	PGN	WIG20
	Grupa Kęty SA	KTY	mWIG40	8	Enea SA	ENA	mWIG40
	Stalprodukt SA	STP	mWIG40		Powszechny Zakład Ubezpieczeń SA	PZU	WIG20
3	Amica SA	AMC	mWIG40	9	Eurocash SA	EUR	WIG20
	CD Projekt SA	CDR	WIG20		Grupa LOTOS SA	LTS	WIG20
	Cyfrowy Polsat SA	CPS	WIG20		Orbis SA	ORB	mWIG40
4	Grupa Azoty SA	ATT	mWIG40		Polska Grupa Energetyczna SA	PGE	WIG20
	Forte SA	FTE	mWIG40		Polski Koncern Naftowy Orlen SA	PKN	WIG20
	ING Bank Śląski SA	ING	mWIG40	10	Budimex SA	BDX	mWIG40
	LC Corp SA	LCC	mWIG40		Ciech SA	CIE	mWIG40
5	Inter Cars SA	CAR	mWIG40	11	Asseco Poland SA	ACP	mWIG40
	Famur SA	FMF	mWIG40		Boryszew SA	BRS	mWIG40
	KGHM Polska Miedź SA	KGH	WIG20		Globe Trade Centre SA	GTC	mWIG40
	mBank SA	MBK	WIG20		Netia SA	NET	mWIG40
6	Pfleiderer Group SA	PFL	mWIG40	12	Bank Pekao SA	PEO	WIG20
	PKO Bank Polski SA	PKO	WIG20		Polimex-Mostostal SA	PXM	mWIG40
	Sanok Rubber Company SA	SNK	mWIG40				

Source: own calculations.