Equilibrium
Quarterly Journal of Economics and Economic Policy
Volume 13 Issue 3 September 2018
p-ISSN 1689-765X, e-ISSN 2353-3293
www.economic-policy.pl

ORIGINAL PAPER

Citation: Chlebus, M. (2018). One-day-ahead forecast of state of turbulence based on today's economic situation. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 13*(3), 357–389. doi: 10.24136/eq.2018.018

Contact: mchlebus@wne.uw.edu.pl, Department of Quantitative Finance, Faculty of Economic Sciences, University of Warsaw, Długa 44/50, 00-241 Warszawa
Received: 6 September 2017; Revised: 28 May 2018; Accepted: 15 June 2018

Marcin Chlebus
University of Warsaw, Poland

One-day-ahead forecast of state of turbulence based on today's economic situation

JEL Classification: *C53; C58; G17*

Keywords: *forecasting; state of turbulence; regime switching; risk management; risk measure; market risk*

Abstract
Research background: In the literature little discussion was made about predicting state of time series in daily manner. The ability to recognize the state of a time series gives, for example, an opportunity to measure the level of risk in a state of tranquility and a state of turbulence independently, which can provide more accurate measurements of the market risk in a financial institution.
Purpose of the article: The aim of article is to find an appropriate tools to predict, based on today's economic situation, the state, in which time series of financial data will be tomorrow.
Methods: This paper proposes an approach to predict states (states of tranquillity and turbulence) for a current portfolio in a one-day horizon. The prediction is made using 3 different models for a binary variable (Logit, Probit, Cloglog), 4 definitions of a dependent variable (1%, 5%, 10%, 20% of worst realization of returns), 3 sets of independent variables (untransformed data, PCA analysis and factor analysis). Additionally, an optimal cut-off point analysis is performed. The evaluation of the models was based on the LR test, Hosmer-Lemeshow test, Gini coefficient analysis and CROC criterion based on the ROC curve. The analyses were performed for 43 individual shares and 5 portfolios of shares quoted on the Warsaw Stock Exchange. The study has been conducted for the period from 1 January 2006 to 31 January 2012.
Findings & Value added: Six combinations of assumptions have been chosen as appropriate (any model for a binary variable, the dependent variable defined as 5% or 10% of worst realization of returns, untransformed data, 5% or 10% cut-off point respectively). Models

built on these assumptions meet all the formal requirements and have a high predictive and discriminant ability to one-day-ahead forecast of state of turbulence based on today's economic situation.

Introduction

In this study, a family of models to predict the state of turbulence for financial time series data have been proposed. The state of turbulence is potentially the most risky and uncertain period. During this state, financial institutions may be exposed to much higher risk than usually and loss much more than expected. Therefore, the ability to recognize a state of turbulence is of crucial interest for financial institutions.

Predictions of a state of turbulence may be used in financial institutions in many ways. They can support the risk management process — for example — by generating a trigger that imposes stricter control processes or increases capital to cover extraordinary losses. The state of turbulence models may also be included in the measurement of market risk in a financial institution. The ability to recognize the state of a time series gives an opportunity to measure the level of risk in a state of tranquility and in a state of turbulence independently, which can provide more accurate measurements of the market risk in a financial institution.

The main aim of the proposed models is to predict, based on today's economic situation, the state in which time series of financial data will be tomorrow.

The rest of the article has been prepared as follows: at the beginning the concept of the proposed models is discussed, then a framework and the testing process is presented, and finally, in order to assess the quality of the proposed models, an empirical analysis has been made.

Concept

The concept — proposed in the study — of the model to predict the state of turbulence for financial time series data was inspired by the process of forecasting the state of the economy (forecasting a crisis). According to my knowledge, the topic of forecasting a state of turbulence (defined as a state where a risk of high losses is relatively high) for financial time series hasn’t been discussed yet, so there is no directly related literature that may be analyzed. Therefore, at the beginning, past studies of the methods used for predicting the state of a crisis have been discussed.

Forecasting the state of the economy is a widely discussed topic among macroeconomists. Prediction of a crisis (a negative state of the economy — a state of turbulence) may help to take appropriate actions to avoid it. Therefore, many researchers have attempted to build models that can predict upcoming crises. These models are called EWS (Early Warning Systems) models.

EWS models, based on information from before the crisis, predict the probability of crisis occurrence within a specified period of time. These models are built on historical information. They are based on the assumption that the crises, despite their differences, have a common specificity that allows them to be treated as (from modelling point of view) homogenous. This hypothesis is discussed in Kamin (1999). However, the author compares only three cases, which do not fully reflect the whole scope of the problem. In fact, we can distinguish several types of crises (such as banking crises, currency crises) that are not only different from each other, but also very often are not independent and follow each other. Therefore, researchers usually build the early warning system for one of the crises types (a banking (Barrel *et al.*, 2010) or a currency crisis (Komulainen & Lukkarila, 2003)).

Sometimes, distinction between types of crises is not sufficient enough. In order to obtain greater homogeneity, the analysis is limited to certain type of economies (such as emerging economies (Komulainen & Lukkarila, 2003). Definitions and classifications of crises do not matter from the perspective of building a model predicting the state of turbulence, since the proposed model is built for a much shorter time horizon. The state of turbulence should be treated as a period of increased risk rather than a period of crisis. Nevertheless, it is important to ensure a proper homogeneity in defining a state of turbulence. Only then is it possible to effectively predict this period.

Despite the difference in the time horizon between classical EWS models and the proposed models, it is possible to use the methodology from EWS models to build the state of turbulence models. The types of models most commonly used to predict a crisis are: signaling models and logistic regression models. These models are used, inter alia, by Kaminsky *et al.* (1998), Beckmann *et al.* (2006), Davis and Karim (2008) and Barrel *et al.* (2010).

Based on the analysis of past studies, there is no conclusion as to which approach is better. Each has its pros and cons, which, depending on the circumstances, make them more or less useful. It can be assumed that the Logit model is better when considering the less specific problems, when the most important thing is to capture the general relation between the occurrence of the crises and the variables under consideration. It means that models of this type are better suited for the analysis of global crises. In contrast, signaling models are better when the problem is considered to be more specific (national crises) (Davis & Karim, 2008).

The aim of the proposed model is to provide a universal tool for predicting the state of turbulence, therefore, the more appropriate model would be the logistic regression model (the Logit model). The most often mentioned disadvantages of the Logit model are its inability to determine the relative quality of a single variable and strength of its impact on the probability of the crisis, as well as, the difficulties in defining precise limits that indicate unusual values for a particular variable (Kaminsky *et al.*, 1998). None of these disadvantages are critical for the model of predicting the state of turbulence, as the only information needed is a prediction on the state conditional on a whole data set of independent variables. In addition, the approach based on the Logit model is more objective than the signalling approach, since the choice of the significance of each variable is independent from the researcher.

While discussing EWS models based on the Logit approach, one of the key problems associated with this type of forecasting models should be described. The results obtained from the Logit model are probabilities of the crisis occurrence (obtained from a transformation of the logarithm of the odds ratio). This means that the model does not predict the crisis, but the probability of its occurrence. In order to obtain a crisis prediction from the probability of a crisis (the value from the interval [0,1]) a threshold should be specified, above which it is considered that the model predicts the crisis. The higher the threshold, the less periods of crisis are going to be predicted. It should increase the share of correctly predicted periods of a crisis and at the same time limit the number of periods falsely predicted as a crisis. On the other hand, the higher the threshold, the more periods of crisis will be considered as periods of tranquility. The choice of the threshold determines, at the same time, the level of the type I and type II errors of the model's predictions. For this reason the choice of the optimal threshold is a very important element in the construction of an EWS model. The level of the threshold should be selected taking into account the high cost of false signals and even higher costs of a crisis (Bussiere & Fratzscher, 2008). Selecting the optimal cut-off point is a problem, which is independent of a time horizon, so it might also be analyzed in the model predicting the state of turbulence.

Classical EWS models are designed to predict a crisis, therefore, the most common perspective of the analysis is annual or biennial. The proposed model is built to predict tomorrow's state of a series of financial data using information about today's economic situation. The difference in the time horizon means that variables usually used in EWS models are not appropriate. Information about the possible data to use in the model can be taken from the studies described by Kim *et al.* (2004) and Oh *et al.* (2006). In these studies, the authors postulate that modern crises are too dynamic to use quarterly data (or less frequent), so EWS models should be built based on daily data. According to them, the best measure of determining the volatility of the economy is the stock index. Additionally, they propose to take into account short-term interest rates and exchange rates. As well as each variable mentioned, its rates of return, moving averages and moving variances might also be included. The proposed set of data seems to be adequate for modelling the state of turbulence.

The studies on daily EWS models are less useful in defining the state of turbulence. This is due to the difference in specificity of modelling the state of turbulence for the economy and for financial time series. In daily EWS models, definition of the state of turbulence refers to the general state of the economy. The state of turbulence for financial data should be more specific and relevant to an analyzed portfolio (asset). It should identify the periods in which the situation of the economy (described by the independent variables) indicates a state of turbulence for the next day.

From a risk management perspective, the most problematic are periods in which losses are the most severe. A market risk management system should be able to identify those periods to protect banks from their consequences. Therefore, it seems reasonable to assume that the state of turbulence for financial data should consist of periods with the highest losses. In this case, it should be possible to predict the periods of most risk and take appropriate steps to protect banks from exceptional losses. For example the state of turbulence forecasting model might be used as a part of Value at Risk model, allowing, in a market risk measurement, to take into account the fact that the some periods are riskier than others.

The above analysis allows to define the basic framework for the state of turbulence forecasting model. Based on the EWS models studies it is possible to choose a possible model to apply (Logit) and the set of independent variables. According to it, it is also important to preserve the homogeneity of the analyzed phenomenon and the choice of the optimal threshold. In addition, the definition of the dependent variable in the model has been proposed as a group of periods with the most severe losses. Discussed topics do not take into account the whole complexity of the framework, but set a starting point, which will be developed to construct an appropriate model for predicting the state of turbulence.

Research methodology

Dependent variable definition

The purpose of the proposed model is to predict the state of turbulence. In order to provide adequate forecasts, first a state of turbulence has to be properly defined. For an analysis of a portfolio (individual asset) the model, as a state of turbulence, should predict the worst cases from the perspective of risk management. The worst case should be connected with the most severe declines in price (returns) of the analyzed portfolio. As the change in the price is non-standardized (it depends on the current value of the assets), a more universal variable is the rate of return of a portfolio (asset). In the study, four different binary dependent variables have been considered: the dependent variable is equal to 1 for 1%, 5%, 10% or 20% of the lowest rates of return and 0 otherwise (dependent variable P1, P5, P10 and P20 respectively).

Dependent variable distribution

According to the conclusions from the previous section, the dependent variable should be a binary variable equal to 1 for the periods defined as periods of turbulence, and 0 for the periods defined as periods of tranquility. For the sake of universality, it is suggested in the literature to use the Logit model. Models for binary dependent variables assume the existence of an unobserved continuous variable, which comes from a specific distribution (in this case, logistic), but only results which takes one of two values (0 or 1) are observed. On the basis of the observed results, the relationship between an unobservable dependent variable and observable independent variables is estimated. The Logit model can be defined as follows (Allison, 2005):

$$y^* = \beta * X + \varepsilon \tag{1}$$

$$y_i = \begin{cases} 1 & if\ y_i^* > 0 \\ 0 & otherwise \end{cases} \tag{2}$$

where y_i^* is a vector of latent variables, β is a vector of parameters, X is a matrix of independent variables, ε is a vector of random error from logistic distribution with parameters equal to $(0;1)$ and $y_i(y_i^*)$ is an observable (latent) result of the phenomenon for i-th observation.

The Logit model assumes that a random error (and therefore unobservable dependent variable) comes from the logistic distribution. Another popular assumption is that a random error comes from a normal distribution. The model with the normal distribution assumption is called the Probit model (Allison, 2005).

The Logit and the Probit models are the most popular models for binary dependent variables. There are also other models for binary dependent variables, which usually allow to overcome problems which the Logit and the Probit models are not able to overcome. Frequently, an important issue in binary variable modelling is the lack of a balance between events and non-events. This is because the observed phenomenon is rare in its nature (e.g., wars, crises). In this case, the Logit and the Probit models may not be adequate (King & Langche, 2001). There are many possible solutions to this problem. The basic model, which takes into account the unbalanced distribution of the dependent variable is the Cloglog model (Complementary log-log) (Allison, 2005). The Scobit model (Nagler, 1994) can also be used or some amendments to the basic models might be performed (King & Langche, 2001). Inclusion of these models or corrections is useful in situations where basic models have a poor quality and it is necessary to search for better solutions. In the study, due to the imbalance of dependent variable, it has been decided to take into account the Cloglog model. The Cloglog model can be defined as the previous two models with the exception that the random error comes from the Gompertz distribution (Allison, 2005). The choice of these three distributions should be sufficient to obtain adequate forecasts of the state of turbulence.

For the above models (Logit, Probit and Cloglog) probability of the state of turbulence is forecasted. The purpose of the model is to predict the state of turbulence instead of the probability of its occurrence. Therefore, the expected result is the value of 1 when the model predicts the state of turbulence and 0 when the model predicts the state of tranquility. To convert a probability (set in the range [0,1]) to a binary forecast of the state, a cut-off point must be defined above which the state of turbulence is forecasted and below which the state of tranquility is forecasted. The study assumed that the optimal cut-off point will be sought with an accuracy of 0.01. Therefore, the analysis of the optimal cut-off point involves cut-off points belonging to the following set:

$$cp_i = 0.01 * i, for\ i = 1, 2, \dots, 100 \tag{3}$$

where cp_i is an i-th cut-off point. Selection of the optimal cut-off point has been carried out on the CROC criterion, which is described later on.

Independent variable data set

The state of turbulence model is based on the assumption that tomorrow's financial data time series state depends on the present state of the economy. The dependent variable is defined as a certain percentage of the worst realization of the return on a portfolio (asset). In order to describe the above relationship, a set of independent variables should appropriately reflect the current market situation. The set of variables that should properly describe the current state of the economy is proposed in Oh *et al.*(2006) and Kim *et al.* (2008). According to the studies, the current state of the economy — on a daily basis — can be determined by variables describing the three key elements of the economy:

1. The situation on the stock market — current values of stock indices;
2. The situation on the currency market — current values of exchange rates;
3. The situation on the interest rates market — current values of short — or mid-term interest rates.

Based on the assumptions presented by the Authors, and by using knowledge of the Polish market characteristics the following variables has been included in the study:

1. Stock indices:
   a. The index of companies listed on the Warsaw Stock Exchange Index (WIG);
   b. The index of the 20 largest companies listed on the Warsaw Stock Exchange (WIG20).
2. Exchange rates:
   a. The Euro to Polish zloty exchange rate (EUR/PLN);
   b. The U.S. dollar to Polish zloty exchange rate (USD/PLN);
   c. The Swiss franc to Polish zloty exchange rate (CHF/PLN).
3. Interest rates:
   a. The overnight WIBOR interest rate (ON);
   b. The 3-month WIBOR interest rate (3M).

Application of these variables only in an untransformed form may not represent the full dynamics of the current economic situation. Therefore, the model should take into account transformations of the aforementioned variables. Using the suggestions made by Kim *et al.*(2008), for each of the variables investigated in this study it was decided to include the following values:

1. The logarithms of rates of return;
2. The 15-day moving average of prices and rates of return;
3. The 15-days moving variances of prices and rates of return.

The choice of the 15-day period for the moving values results from the trade-off between the dynamics of independent variable changes and the stability of the relationship between a dependent and independent variables. Taking into account all the transformations of the aforementioned variables, 35 potential independent variables has been considered (5 different values for 7 variables).

It is worth noting that some of them may be highly collinear, which can have a negative impact on the quality of the model. Collinearity may cause the suggested parameters to not properly capture the dependencies between independent variables and the dependent variable. These parameters may additionally reflect (part of) the relationship between the collinear independent variables and the dependent variable. It might be an important issue when the objective of the model is to describe the influence of a single independent variable on the dependent variable. It is less important when the objective of the model is to describe the cumulative effect of a set of independent variables on the dependent variable (forecast the dependent variable). Then, even if individual parameters are biased because of collinearity, the cumulative effect of a set of independent variables on the dependent variable may remain appropriate.

There is no one method of solving the collinearity problem. The simplest approach is to leave all variables in the model and take it into account during the interpretation of the results. There are other, more sophisticated, solutions suggested in other studies. One possibility is to exclude from the set of independent variables those variables that cause the problem of collinearity. This solution may cause a problem with the correct specification of the model. Another possible solution is to use methods of reducing the dimensions of the matrix (the PCA or the factor analysis). These approaches, on the one hand, limit the set of independent variables analyzed, and on the other hand, solve the problem of collinearity, as calculated components or factors are orthogonal to each other (Greene, 2003). In this study it was decided to consider three possible solutions to the problem of collinearity: leaving all the variables in the model (untransformed data set); taking the principal components obtained from the PCA (with orthogonal rotation using varimax methods) as an independent variables data set; taking factors obtained from factor analysis (with orthogonal rotation using varimax methods) as an independent variables data set.

It has turned out that variables based on the WIG and the WIG20 indices are almost exactly collinear and for this reason it has been decided to exclude the variables based on the WIG index and leave only the variables based on the WIG20 index. In the final set of independent variables, 30 independent variables has been included.

The proposed set of independent variables has been designed to be universal (to describe the current state of the economy). The opposite assumption has been made in the case of the definition of dependent variable, which is specific to each portfolio (asset). The dependent variable is defined based on the characteristics of a specific portfolio (asset). Therefore, it should be taken into account that the current state of the economy (represented by a set of independent variables) may have a different impact on the state of turbulence forecasts for different portfolios (assets). This should be taken into account by specifying the parameters for each portfolio (asset) separately. This approach caters for the universality arising from the same set of independent variables and the specificity related to the definition of the dependent variable. Both properties are preserved regardless of whether the model is built on untransformed data, the PCA data, or the factor analysis data.

Model testing

The proposed models for predicting the state of turbulence may differ from one another in four dimensions: random error and latent variable distribution (Logit, Probit, Cloglog), definition of the dependent variable (P1, P5, P10, P20), set of independent variables (untransformed data, the PCA data, the factor analysis data) and cut-off point (100 possible values).

This diversity means that 3600 different combinations of assumptions need to be considered. Therefore, the testing process should be wide enough to find out imperfections of as many combinations of assumptions as possible. For this purpose, the testing process involve four different analyses: the Hosmer-Lemeshow goodness-of-fit test (null hypothesis: model is well fitted to data), the LR test for independent variables insignificance (null hypothesis: all variables in the model are jointly insignificant), Gini coefficient analysis (discriminant ability of the model) and CROC criterion (prognostic ability of the model).

The proposed set of analyses is designed to assess the appropriateness of assumptions, both from the formal perspective (Hosmer-Lemeshow test, LR test), and from the performance perspective (Gini coefficient, CROC criterion — distance between a point on the ROC curve and the ideal point). The Hosmer-Lemeshow and the LR tests are typically used to assess the quality of models for binary variables (Hosmer & Lemeshow, 2000). The discriminant and predictive ability measures are not always used for such assessments, however, they are gaining in popularity. They are often used for assessment of the probability of default models (credit risk models).

The Gini coefficient is used to analyze the discriminant ability of the model. It allows to assess how the model forecasts separate distributions of successes and failures. An effective model should assign a high probability of success for observations, which are in fact successes and low probability of success for observations, which are in fact failures. The Gini coefficient is calculated based on the CAP curve (Cumulative Accuracy Profile), which is a graphical illustration of the distribution of the probability of success conditional on observed success in relation to the unconditional distribution of the probability of success (BCBS, 2005). Examples of CAP curves are shown on Figure 1.

The figure shows three possible shapes of the CAP curve. The curve for the model with full discrimination (dashed line), the curve for an example model (solid line) and the curve for the model with lack of discrimination (dotted line). Assuming that the model is not worse than the model with lack of discrimination, its CAP curve may be located between two extreme CAP curves (full and lack of discrimination). The higher discriminant ability that a model has, the closer the CAP curve for the model is to the curve for full discrimination.

The Gini coefficient is calculated based on the relationship between the three CAP curves presented on Figure 1. It is equal to the ratio of the area between the CAP curves for the analyzed model and the model without discrimination, and the area between the CAP curves for the model with full discrimination and the model without discrimination, which can be written as follows:

$$GINI = \frac{a_m}{a_m + a_p} \tag{4}$$

where a_m is an area between CAP curves for the analyzed model and a model without discrimination, a_p is an area between a CAP curves for the model with full discrimination and a model without discrimination.

Values for the Gini coefficient are in the interval [-1,1]. The closer the coefficient value is to 1, the stronger is the model's discriminant ability (Tasche, 2008). Unfortunately there are no clear critical values (thresholds) for this measure, which would separate high-quality models from low-quality ones. Thresholds are determined based on experience and are relative (Anderson, 2007). The necessary minimum is that the value of the Gini coefficient is greater than 0, which means that the discriminant ability of the model is greater than in the case of lack of discrimination.

The CROC criterion based on ROC curve analysis may be used for a predictive ability assessment. The ROC curve illustrates the relationship between two independent distributions. It is determined by the relation between the distributions of the success conditional on observed successes and observed failures. Therefore, the ROC curve consists of a set of points $(F_L(s), F_S(s))$, where s comes from the interval $[0,1]$, $F_L(.)$ denotes the cumulative distribution function of the probability of success conditional on observed failures, and $F_S(.)$ denotes the cumulative distribution function of the probability of success conditional on observed success. The ROC curve (usually) is described in terms of two measures: the sensitivity and the specificity. The sensitivity measures the ability of the model to correctly predict success. The specificity measures the ability of the model to correctly predict failures. These measures are calculated for binary variables, which means that in order to calculate them, it is necessary to transform probability of success into prediction of success or failure. The sensitivity and the specificity may be calculated for each cut-off point s (Tasche, 2008).

Comparing models using these two measures independently is usually inconclusive, because one of the models may be better in terms of sensitivity and the other in terms of specificity. In order to obtain a measure that allows unambiguous comparison of the models, weighted indexes are created. Weights for the sensitivity and the specificity are associated with the relative costs of incorrect success and failure forecasts (Steyerberg *et al.*, 2011). The most common measure of predictive ability calculated on the basis of the specificity and the sensitivity of a model is the Youden index (Youden, 1950). It assumes that the trade-off between the sensitivity and the specificity is linear with a scaling factor equal to one. This implies that at a certain sum of the sensitivity and the specificity, the extreme solutions are equally good (the sensitivity or the specificity is equal to 0) as moderate solutions (the sensitivity and the specificity are different from 0). To involve a preference of moderate solutions over extreme solutions, instead of the Youden index, the CROC criterion may be used. According to this criterion, the best model is the one that minimizes the distance between the ideal point and a point on the ROC curve (depends on a cut-off point). The ideal point is the point in the upper left corner of the ROC curve graph. At this point, the sensitivity and the specificity are equal to 1 (model predicts success for all observed successes and failure for all observed failures). The CROC criterion value can be calculated as follows:

$$CROC = \sqrt{\left(\frac{FN}{TP+FN}\right)^2 + \left(\frac{FP}{TN+FP}\right)^2} \tag{5}$$

where TP is the true positive rate, TN is the true negative rate, FP is the false positive rate and FN is the false negative rate.

In this case, the isocosts curves (curves defining models of the same quality) have the shape of a quarter circle with center at the ideal point. This measure prefers models with a smaller value of a root of a sum of squares of type I and type II errors (incorrect predictions of successes and incorrect predictions of failures). In this case, the best solution, assuming a certain sum of the sensitivity and the specificity, is the middle solution (the sensitivity and the specificity are equal), and the extreme solutions are the least attractive (Powers, 2011).

Both measures have their optimal solutions in the ideal point, with the difference that the optimal value of the Youden index is equal to 1, and the optimal value of the CROC criterion is equal to 0. To assess the predictive ability in the study, the CROC criterion has been selected, the reason being that, with a certain sum of the specificity and the sensitivity, moderate solutions are preferred over extreme solutions. From the state of turbulence predictions perspective, a model that predicts only a state of turbulence or only a state of tranquility should be less attractive than the model which predicts both states. For the CROC criterion, as for the Gini coefficient, no limits are defined which distinguish high-quality from poor-quality models. The CROC criterion is used to provide relative comparisons of the models.

The testing process presented above should allow to choose the best foundation for the state of turbulence predicting model. The test procedure consists of two stages. First, the formal tests should be performed. Then, discrimination and predictive ability should be assessed. Formal tests provide information as to whether the set of assumptions is good enough to be used. The Gini coefficient and the CROC criterion allow to compare models with different assumptions. Finally, based on the results the best possible set of assumptions for the state of turbulence predicting model may be selected.

Empirical results

Individual asset analysis results

The empirical research on the state of turbulence predicting models has been divided into two parts. First, the analysis has been made for a series of rates of return on individual assets, then the analysis has been made on a series of rates of return on portfolios consisting of 10 randomly selected assets. The analysis has been made for shares listed on the Warsaw Stock

Exchange. For the individual assets, in-sample analysis has been performed. For the portfolios, in-sample analysis and out-of-sample analysis has been performed.

Three sets of independent variables have been considered: the untransformed data set, the principal components obtained from the PCA (with orthogonal rotation using varimax approach) analysis and the factors (with orthogonal rotation using varimax approach obtained from the factor analysis. Taking into account that the set of untransformed independent variables, for each of the assets and the portfolio, is the same and all of the models have been constructed based on the same data from the same period, both the untransformed data set itself and the results of the observation matrix reduction methods (the PCA and the factor analysis) are the same for all assets and portfolios.

The PCA analysis and the factor analysis have been performed on all 30 independent variables. The principal components have been selected with respect to the Scree plot and the Kaiser criterion. It has been decided to select six principal components. The factors have been selected based on the Scree plot and the Variance Explanations criterion. According to the results, five factors have been selected.

Each of three independent variables data sets (untransformed data, principal components and factors) have been used to estimate the state of turbulence models for each of the assets and for each of the portfolios independently.

Analysis for an individual asset has been conducted for 43 different shares listed on Warsaw Stock Exchange. The shares have been chosen randomly. The only condition imposed was that shares had to be listed on the Warsaw Stock Exchange since at least January 2006. The study covered the period from 1 January 2006 to 31 January 2012. The companies whose shares were included in the study are presented in Table 1. All data was downloaded from the *stooq.pl* web service.

A wide range of assets have been taken into consideration which should allow detail verification of the correctness of the analyzed sets of assumptions. It is worth restating that each set of assumptions for the state of turbulence predicting models has been tested on 43 different dependent variables.

An analysis of the validity of the assumptions made has been performed in accordance with the testing process described above. The results are presented in an aggregated manner — as the average score for all 43 assets. In this study, in-sample analysis has been performed. This analysis is made for the observations on which the model has been estimated.

Formal tests

The testing process is begun by performing the Hosmer-Lemeshow test. Table 2 presents the percentage of cases for which the null hypothesis in Hosmer-Lemeshow test has not been rejected in models with various theoretical distributions, definitions of the dependent variable and data types.

For all analyzed significance levels (10%, 5% and 1%) this test does not significantly prefer any of three theoretical distributions. However, in all cases the best results have been achieved for the Probit models. The results for all the theoretical distributions are good enough to consider the possibility of using each of them in the state of turbulence models.

For dependent variables P5, P10, and P20 the best results are obtained from models built on untransformed data (Table 2) omits results for the principal components and the factors for these variables because the results were much worse than for untransformed data and are not worth considering). For the P1 variable all independent variables data sets are of the same quality (actually the untransformed data has the worst results). From analysis of models built on the untransformed data for the different definitions of the dependent variable it may be seen that the models for the P1 variable achieve relatively the worst results. Actually, only for the P1 dependent variable the percentage when the null hypothesis has not been rejected is much smaller than the expected result.

Results for the Hosmer-Lemeshow test indicate that in the state of turbulence models, all considered theoretical distributions of a random error may be used. In addition, for the P5, P10 and P20 variables, models achieve the best results when using the untransformed data. The only exception is the P1 dependent variable, for which slightly better results have been achieved for the models based on the data from the factor analysis. The Hosmer-Lemeshow test does not disqualify any of the dependent variables, although the poorest results have been obtained for the variable P1.

The next step in the testing process is the analysis of the LR test results. The purpose of this test is to test the total irrelevance of the impact of the independent variables on the dependent variable. The table with results of the LR test shows the percentage of cases in which the null hypothesis of the LR test has been rejected. Results obtained for the LR test are presented in Table 3. Presented results are for models estimated on the untransformed data. Results for models estimated on the data from the PCA analysis and the factor analysis were significantly worse and have been omitted.

LR test results for each definition of the dependent variable indicate that independent variables are much more likely to be irrelevant for the P1 dependent variable than for the others. Among the P5, P10 and P20 dependent variables, the P20 variable is relatively the worst. It has also turned out that the significance of the impact in the cases of the P5 and the P10 dependent variables is similar. In addition, similar results of the LR test have been obtained for all random error distributions. All models (Logit, Probit and Cloglog) are equally good with respect to this test.

Results of the formal tests (the Hosmer-Lemeshow test and the LR test) can be summarized in the following points:

The best results are obtained for the state of turbulence models built on a set of untransformed data.

On the basis of the formal tests analyzed, distributions of random errors are equally good and might be used in the state of turbulence models.

Formal tests do not unequivocally reject any of the dependent variables, although for the LR test it can be observed that worst results are obtained for the P1 variable.

The next step is to measure the discriminant and the predictive ability of the model analysis. For this purpose, the Gini coefficient and the CROC criterion have been used.

Gini coefficient analysis

The Gini coefficient is a measure of the discriminant ability of a model. It determines how well the predicted probability of the state of turbulence separates the distribution of rates of return from the state of turbulence from the distribution of rates of return from the state of tranquility. Results of the Gini coefficient analysis are presented in Table 4. In the table the average values of the Gini coefficient are presented for the state of turbulence models based on the untransformed data. Again, results for the PCA data and for the factor analysis data have been omitted. This is due to the fact that the assessment of the discriminant ability of the models based on different types of data clearly indicates that the discriminant ability of the models based on the untransformed data is significantly higher. On the basis of these results, it may be concluded that the use of principal components or factors instead of untransformed data to build the state of turbulence model worsens its discriminant ability. It can also be stated that, despite the fact that observation matrix reduction methods solve the problem of collinearity, their use may lead to a worsening of the quality of information stored in the data set.

Based on the results shown in the Table 4, it may be concluded that the smaller the area that defines the state of turbulence, the higher the discriminant ability of the model. According to the results obtained, models with the P1dependent variable have the largest discriminant ability, models with the P5 and the P10 dependent variables have moderate discriminant ability and models with P20 dependent variable have definitely the worst discriminant ability. These results might be interpreted as follows: the more extreme realization of returns are, the more similar and significantly different from those in a defined state of tranquility the situation on the market is. On the basis of the market situation, it is easier to recognize more extreme realizations of returns. It means that the expansion of the definition of the state of turbulence increases the noise. However, it is important to remember about a risk that using too narrow definition of the state of turbulence may lead to model over-fitting. It means that the Gini coefficient for the P1dependent variable might be so high not due to the actual relationship between today's situation in the economy and tomorrow's financial time series state, but due to specific relationship for a data set. In this case, it may happen that the model, according to in-sample analysis, works correctly, but in fact, when used in reality would perform much worse. In order to verify the over-fitting issue, out-of-sample analysis should be performed. The description of its construction and its effect on the state of turbulence forecasting models will be presented in the portfolios analysis section.

As for formal tests, the Gini coefficient results are similar and good for all three considered assumptions about the distribution of a random error. The similarity is confirmed for each of the definitions of the dependent variable. Again, in all cases, the average value of the Gini coefficient is the highest for the Probit model, but the differences are very small and do not appear to be significant.

Based on the results obtained during the formal testing and the discriminant analysis, the following conclusions may be stated:

- All theoretical distributions of error terms may be used to build the state of turbulence model — the quality of the Logit, Probit and Cloglog models (ceteris paribus) are good enough and do not differ significantly,
- Models built on the untransformed data achieve significantly better results than the models based on the principal components or the factors. Therefore, for further analysis only the untransformed data will be considered,
- Models in which the dependent variable is defined as 20% of the worst realizations of rates of return, achieve much worse results than models for other dependent variables. For this reason, the P20 dependent variable will be excluded from a further analysis.

CROC criterion analysis

The primary purpose of the proposed models is to provide high-quality state of turbulence forecasts. Therefore, the CROC criterion results are crucial for their evaluation. In this analysis, a set of assumptions has been extended by adding a cut-off point analysis. This is very important, as two identical models which differ only in the cut-off point assumption may have completely different predictive abilities. For the formal tests and the Gini coefficient, a cut-off point value does not matter. These analyses are prior to the step of determining the cut-off point. The CROC criterion allows to consider an extra dimension of a model, it allows to evaluate which cut-off point for the state of turbulence model achieves the best results.

Results for the CROC criterion analysis are presented in Tables 5–7. These tables include the average distance, for different cut-off points, between the point on the ROC curve and the ideal point, for models with differing assumptions. The lower value of the CROC criterion is, the better predictive ability a model has. The tables show results for cut-off points around the optimal cut-off point (the CROC criterion is on average the lowest).

Analyzing values presented in the tables, it can be stated that the average values of the CROC criterion for the optimal cut-off points are significantly lower (better) for the dependent variable P1 (0.185–0.203) than the average values for the variable P5 (0.412–0.418) and the variable P10 (0.482–0.487). Results for the P5 and the P10 variables are very similar and equally good.

Again, the results for the various theoretical distributions of a random error do not differ from each other significantly, although at the optimal cut-off points the best results are always achieved for the Probit model. Considering the average results for the CROC criterion for different theoretical distributions of a random error, it may be stated that the results for the Cloglog, the Logit and the Probit models are basically indistinguishable.

In the CROC criterion results interesting regularity may be seen. On average, the lowest values of this criterion have been achieved when the cut-off point is equal to the percentage that defines the state of turbulence (dependent variable). It means that, on average, for the P1 dependent variable the optimal cut-off point is equal to 0.01, for the P5 variable is equal to 0.05, and for the P10 variable is equal to 0.1.

Based on the CROC criterion analysis, the conclusions may be summed up in the following points:

The smaller the area that defines a dependent variable, the better the predictive ability of a model. The best results have been achieved for the P1 dependent variable, results for the P5 and the P10 variables are slightly worse. Nevertheless, the results for the P5 and the P10 dependent variables do not exclude them from a further analysis.

The CROC criterion does not materially prefer any one of the family of distributions of a random error.

The optimal cut-off point, on average, should be equal to a percentage that defines a dependent variable (the state of turbulence).

The testing process consisted of four elements: the Hosmer-Lemeshow test, the LR test, the discriminant ability analysis based on the Gini coefficient and the predictive ability analysis based on the CROC criterion. The results obtained during the testing process have enabled to reduce 3600 different combinations of assumptions to 9 possible combinations, which provide relatively the best state of turbulence models. The best combinations of the assumptions have been presented in the Table 8.

The results obtained during individual assets analysis have pointed out the 9 sets of assumptions that should define high quality models. The correctness of these conclusions has been verified in the portfolio analysis. The results of the portfolios analysis are presented in the next section.

Portfolio analysis results

In portfolios analysis five portfolios have been examined. Each of them consists of 10 randomly selected assets listed on the Warsaw Stock Exchange. Portfolios are built on the assumption that at any time the share of each asset in the portfolio is fixed and equal to 1/10 of the portfolio value. This is a simplifying assumption, which should not affect the result of the analysis. Same as for the analysis of individual assets, the assets in the portfolios are listed at least 6 years on the Warsaw Stock Exchange. The study has been conducted for the period from 1 January 2006 to 31 January 2012.

In the study, in order to verify the results obtained during the analysis for individual assets, in-sample and out-of-sample analysis has been performed. The out-of-sample analysis enables to check the stability of the results obtained during the in-sample analysis. The out-of-sample analysis is made for observations, which have not been taken into account during estimation. A significant reduction of the forecast quality implies an overfitting issue which generates a high risk that the model may perform significantly more weakly in reality than it is expected to according to in-sample results. In this case, the state of turbulence model instead of describing a universal relationship between the dependent variable and the independent variables, describes the relationship between them resulting from the specificity of the analyzed data set.

The out-of-sample analysis has been performed in accordance with the recursive window approach. The analysis has been carried out for the period between 1 January 2010 and 31 January 2012.

The initial training sample consisted of 1004 observations and the prediction sample consisted of 525 observations. This means that the out-of-sample analysis has been made on a sample of 525 forecasts.

The test procedure, as in models for individual assets, has consisted of four components: the Hosmer-Lemeshow test, the LR test, the Gini coefficient and the CROC criterion analyses. Formal tests, due to their specificity (assess the goodness of fit and form of the model) have been carried out only during the in-sample analysis. The Gini coefficient and the CROC criterion analyses have been carried out both for the in- and out-of-sample analysis.

First, the Hosmer-Lemeshow test (Table 10) was performed. The results of the test have confirmed the conclusions drawn from the analysis for individual assets. For all analyzed portfolios, in most cases, there is no reason to reject the null hypothesis, because of the correctness of the assumed theoretical distribution of a random error (for the significance level equal to 5%). The worst results have been obtained by the Cloglog model with the P1dependent variable. For this model for 3 out of 5 portfolios the percentage of cases in which there was no reason to reject the null hypothesis is lower than the expected 95%. For the Probit model with the P10 dependent variable, the same results have been obtained for two portfolios. It can also be noted that for portfolio 1 and dependent variable P5 none of the considered theoretical distributions reach the expected percentage of cases in which there is no reason to reject the null hypothesis. However, for all described cases results are as expected. Therefore, despite the imperfections, it is difficult to reject any combination of the assumptions for the state of turbulence models.

Then the LR test has been performed (Table 11). It may be concluded that for the assumed level of significance equal to 5%, untransformed independent variables always have significant combined impact on a dependent variable. All combinations of analyzed assumptions, from the perspective of this test, are equally good.

The next stage of the evaluation was the Gini coefficient analysis. This coefficient was calculated both during the in-sample and the out-of-sample analysis. The Gini coefficient values for the in-sample analysis (Table 12) confirm the results obtained during the individual assets analysis. Again, the smaller the area that defines the dependent variable, the greater the discriminant ability of the model. In addition, for each dependent variable the Logit, the Probit and the Cloglog model have a very similar discriminant ability.

The key results have been obtained during the out-of-sample analysis (Table 13). These results indicate that models built on the P10 and the P5 dependent variables do not lose much of their discriminant ability when out-of-sample analysis is performed. In contrast, models with the P1 dependent variable have lost their discriminant ability very significantly. Extremely bad results have been reached for the portfolio number 1, where the Gini coefficient is smaller than 0. This means that the model has a negative discriminant ability. Obtaining such a result suggests a very high instability for the P1 dependent variable.

The final step of the testing process was the CROC criterion analysis. It was performed during the in-sample and the out-of-sample analyses. All values presented below have been calculated with the assumption that the cut-off point is equal to the percentile defining the dependent variable. This means that the cut-off point for the P1variable is equal to 0.01, the cut-off point for the P5 variable is equal to 0.05 and for the P10 variable equal to 0.1.

The CROC criterion results for the in-sample analysis (Table 14) confirm the results obtained for individual assets. The CROC criterion values are significantly lower for the P1dependent variable than for the P5 and the P10 variables. Again, the CROC criterion values for the P5 and the P10 variables are similar and low enough to assume that models with those dependent variables have high predictive quality. For each of the portfolios under consideration, it is possible to find cut-off points with a lower value of the CROC criterion than for cut-off points equal to the percentile which defines the dependent variable. However, the results for the so- defined cut-off points are close to optimal, and each time provide a model with a high predictive ability.

As for the Gini coefficient, the CROC criterion out-of-sample analysis (Table 15) has provided the key results. It has turned out that during the out-of-sample analysis, models built on the P5 and the P10 dependent variables have lost much less from their predictive ability in comparison to models built on the P1 variable. Models with the P1variable have a very poor prognostic ability. For portfolios 1, 2 and 3, the binary models (Logit, Probit and Cloglog) have not once predicted the state of turbulence correctly. This result confirms the conclusions of the Gini coefficient out-of-sample analysis. Models based on the P1 dependent variable are characterized by high instability. This means that models based on the P1variable, although they work well for the sample on which the model has been estimated, its usefulness in the predicting process is much smaller.

Conclusions

During the empirical results analysis a number of possible combinations of assumptions for the state of turbulence models have been examined. It has been studied how different assumptions affect the quality of the model. The assumptions about the theoretical distribution of a random error, the definition of the dependent variable, the types of the independent variables data set and the choice of the cut-off point that defines the state of turbulence have been taken into consideration.

Based on the results of the in-sample analysis for individual assets, it may be stated that the choice of the theoretical distribution of a random error from the normal, the logistic and the Gompertz distribution is irrelevant to the quality of the model (all assumptions are equally good). The results have shown that models with the P20 dependent variable or ones built on the principal component or the factors have much lower quality than the others. It might also be noted that, on average, the optimal cut-off point is equal to the percentage of observations that defines the state of turbulence.

The in-sample analysis of the portfolios has confirmed the previously obtained results. Additionally, the out-of-sample analysis implies that models built on the P1 dependent variable are characterized by a very high instability in its discriminant and predictive ability.

Finally, the six combinations of assumptions, which meet all formal requirements and have a high predictive and discriminant ability, both for the in-sample and the out-of-sample analyses, have been selected. Table 16 shows the possible groups of assumptions which should allow to build a high quality state of turbulence model.

Due to the differences in the horizon of the analysis and significantly different specificity of the state of turbulence model and the EWS models developed by the authors of the studies discussed above, it is difficult to directly compare the obtained results. Nevertheless, it is worth referencing the general conclusions obtained in the study with the conclusions from the EWS model studies.

Firstly, it is worth noting that the results of the study confirm the findings from studies such as Eichengreen *et al.* (1995), Kaminsky *et al.* (1998), Beckmann *et al.* (2006), Davis and Dilurby (2008), Bussiere and Fratzscher (2008) and Barrel *et al.* (2010), in which models for binary dependent variable are regarded as adequate to predict the state of turbulence (the state of the crisis). The choice of the optimal cut-off point has also turned out to be very important as in the study of Fratzscher and Bussiere (2008).

Secondly, the results obtained have also confirmed the conclusions from Oh *et al.* (2006) and Kim *et al.* (2008) about the set of independent variables that can be used to predict the state of turbulence. The inclusion of variables describing the situation on the stock exchange, the foreign exchange and the interest rates markets should provide effective forecasts of the state of turbulence.

It seems that the state of turbulence models constructed in accordance with the scheme described in the study, with one of the six combinations of assumptions selected, can provide high quality forecasts and thus be very useful in risk management in a financial institution. The groups of assumptions specified in the Table 16, not only meet the formal requirements, but are also characterized by a high and stable predictive and discriminant ability. The high forecasts quality leads to the conclusion that the proposed model may be an effective tool for generating signals that trigger stricter control processes or increases in a capital buffer with respect to extraordinary loss expectation. It seems that these models may also be effectively used in measuring the market risk in a financial institution. In this case, its usefulness should be verified on the basis of quality of the risk level forecasts provided by the models used to measure market risk, which would take into account the state of turbulence models. The possibility of using forecasts of the state of turbulence in the measurement of the market risk is an important issue because it could potentially lead to improved estimates of the regulatory capital held for market risk by a financial institution, which may translate into a greater stability of the entire sector. Construction and evaluation of the effectiveness of the market risk measurement models using state of turbulence models is a direction worth developing for the proposed models.

References

Allison, P. (2005). *Logistic regression using SAS: theory and application*. Cary, NC: John Wiley & Sons.

Anderson, R. (2007). *The credit scoring toolkit: theory and practice for retail credit risk management and decision automation.* Oxford University Press.

Beckmann, D., Menkhoff, L., & Sawischlewski, K. (2006). Robust lessons about practical early warning systems. *Journal of Policy Modeling*, *28*(2). doi: 10.1016/j.jpolmod.2005.10.002.

Bussiere, M., & Fratzscher., M. (2008). Low probability, high impact: policy making and extreme events. *Journal of Policy Modeling*, *30*(1). doi: 10.1016/j.jpolmod.2007.03.007.

Bussiere, M., & Fratzscher., M. (2006). Towards a new early warning system of financial crises. *Journal of International Money and Finance*, *25*(6). doi: 10.1016/j.jimonfin.2006.07.007.

Davis, P., & Karim, D. (2008). Comparing early warning systems for banking crises. *Journal of Financial Stability*, *4*(2). doi: 10.1016/j.jfs.2007.12.004.

Demirguc-Kunt, A., & Detragiache, E. (1998). The determinants of banking crises in developing and developed countries. *IMF Staff Papers*, *45*(1). doi: 10.5089/9781451947175.001.

Demirguc-Kunt, A., & Detragiache, E. (2000). Monitoring banking sector fragility: a multivariate Logit approach. *World Bank Economic Review*, *14*(2). doi: 10.5089/9781451947175.001.

Eichengreen, B., Rose, A., Wyplosz, C., Dumas, B., & Weber, A. (1995). Exchange market mayhem: the antecedents and aftermath of speculative attacks. *Economic Policy*, *10*(21). doi: 10.2307/1344591.

Greene, W. (2003). *Econometric analysis*. Prentice Hall.

Hosmer, D., & Lemeshow, S. (2000). *Applied logistic regression*. John Wiley & Sons.

Kamin, S. (1999). The current international financial crisis: how much is new? *Board of Governors of the Federal Reserve System International Finance Working Paper*, *636*. doi: 10.2139/ssrn.171714.

Kaminsky, G., Lizondo, S., & Reinhart, C. (1998). Leading indicators of currency crises. *IMF Staff Papers*, *45*(1). doi: 10.5089/9781451955866.001.

Kaminsky, G. (1998). Currency and banking crises: the early warnings of distress. *Board of Governors of the Federal Reserve System International Finance Working Paper*, *629*. doi: 10.5089/9781451858938.001.

Kim, H. J. (2008). Common factor analysis versus principal component analysis: choice for symptom cluster research. *Asian Nursing Research*, *2*(1). doi: 10.1016/s1976-1317(08)60025-0.

Kim, T. Y., Hwang, C., & Lee, J. (2004). Korean economic condition indicator using a neural network trained on the 1997 crisis. *Journal of Data Science*, *2*.

King, G., & Langche, Z. (2001). Logistic regression in rare events data. *Political Analysis*, *9*. doi: 10.1093/oxfordjournals.pan.a004868.

Komulainen, T., & Lukkarila, J. (2003). What drives financial crises in emerging markets? *BOFIT Discussion Papers*, *5*. doi: 10.2139/ssrn.1015459.

Nagler, J. (1994). Scobit: an alternative estimator to Logit and Probit. *American Journal of Political Science, 3*8(1). doi: 10.2307/2111343.

Oh, K. J., Kim, T. Y., & Kim C. (2006). An early warning system for detection of financial crisis using financial market volatility. *Expert Systems*, *23*. doi: 10.1111/j.1468-0394.2006.00326.x. `

Powers, D. (2011). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal Of Machine Learning Technologies*, *2*(1).

Steyerberg, E., Van Calster, B., & Pencina, M. (2011). Performance measures for prediction models and markers: evaluation of predictions and classifications. *Revista Espanola de Cardiologia (English Edition)*, *64*(9). doi: 10.1016/j.rec.2011.05.004.

Studies on the validation of internal rating systems. Basel Committee on Banking Supervision: Basel. Retrieved from http://www.bis.org/publ/bcbs_wp14.pdf. (08.08.2016).

Tasche, D. (2008). Validation of internal rating systems and PD estimates. In G. Christodoulakis & S. Satchell (Eds.). *The analytics of risk model validation*. Academic Press. doi: 10.1016/b978-0-7506-8158-2.x5001-x.

Youden, W. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1). doi: 10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3.

Annex

Table 1. Companies, which shares were included in the study for individual assets

NO	Company name	NO	Company name	NO	Company name
1	ASSECO POLAND S.A.	16	FERRUM S.A.	31	PROJPRZEM S.A.
2	AMPLI S.A	17	FAMUR S.A.	32	OPAKOWANIA PLAST-BOX S.A.
3	BETACOM S.A.	18	INSTAL KRAKÓW S.A.	33	POLNORD S.A.
4	BRE BANK S.A.	19	KCI S.A.	34	SOPHARMA AD
5	CERAMIKA NOWA GALA S.A.	20	KGHM S.A.	35	STALEXPORT AUTOSTRADY S.A.
6	COGNOR S.A.	21	KOGENERACJA S.A.	36	SWISSMED CENTRUM ZDROWIA S.A.
7	CENTROZAP S.A.	22	LPP S.A.	37	TELL S.A.
8	DOM DEVELOPMENT S.A.	23	MCLOGIC S.A.	38	TRION S.A.
9	ECHO INVESTMENT S.A.	24	MENNICA POLSKA S.A.	39	TELEKOMUNIKACJA POLSKA S.A.
10	EFEKT S.A.	25	MOSTOSTAL PŁOCK S.A.	40	VISTULA GROUP S.A.
11	ELEKTRO BUDOWA S.A.	26	MOSTOSTAL WARSZAWA S.A.	41	WASKO S.A.
12	ELZAB S.A.	27	MOSTOSTAL-EXPORT S.A.	42	WILBO S.A.
13	ENERGOMONTAŻ-POŁUDNIE S.A.	28	MOSTOSTAL ZABRZE - HOLDING S.A.	43	ŻYWIEC S.A.
14	FAM GK S.A.	29	MUZA S.A.		
15	FARMACOL S.A.	30	NORDEA BP S.A.		

Table 2. The Hosmer-Lemeshow test results for the individual assets analysis

MODEL	DEP. VAR.	DATA TYPE	SIGNIFICANCE LEVEL		
			1%	5%	10%
P1	FA	Cloglog	95.2%	95.2%	92.9%
P1	FA	Logit	95.2%	95.2%	92.9%
P1	FA	Probit	97.6%	95.2%	90.5%
P1	Untransformed	Cloglog	93.0%	86.0%	83.7%
P1	Untransformed	Logit	90.9%	88.6%	88.6%
P1	Untransformed	Probit	95.5%	95.5%	93.2%
P1	PCA	Cloglog	93.0%	93.0%	90.7%
P1	PCA	Logit	93.0%	93.0%	88.4%
P1	PCA	Probit	95.3%	90.7%	86.0%

Table 2. Continued

MODEL	DEP. VAR.	DATA TYPE	SIGNIFICANCE LEVEL		
			1%	5%	10%
P5	Untransformed	Cloglog	97.7%	93.2%	84.1%
P5	Untransformed	Logit	97.7%	95.5%	84.1%
P5	Untransformed	Probit	100.0%	95.5%	90.9%
P10	Untransformed	Cloglog	100.0%	97.7%	95.5%
P10	Untransformed	Logit	100.0%	100.0%	100.0%
P10	Untransformed	Probit	100.0%	97.7%	97.7%
P20	Untransformed	Cloglog	100.0%	95.5%	93.2%
P20	Untransformed	Logit	100.0%	95.5%	93.2%
P20	Untransformed	Probit	100.0%	97.7%	95.5%

The table shows the percentage of cases in which there is no reason to reject the null hypothesis in the Hosmer-Lemeshow test for the state of turbulence models with different assumptions about the distribution of a random error, a set of independent variables and the dependent variable definition.

Table 3. The LR test results for the individual assets analysis

DEP. VAR.	DATA TYPE	MODEL	1%	5%	10%
P1	Untransformed	Cloglog	58.1%	72.1%	79.1%
P1	Untransformed	Logit	54.5%	70.5%	79.5%
P1	Untransformed	Probit	54.5%	68.2%	79.5%
P5	Untransformed	Cloglog	72.7%	86.4%	90.9%
P5	Untransformed	Logit	72.7%	88.6%	90.9%
P5	Untransformed	Probit	72.7%	88.6%	90.9%
P10	Untransformed	Cloglog	75.0%	84.1%	90.9%
P10	Untransformed	Logit	75.0%	86.4%	90.9%
P10	Untransformed	Probit	77.3%	88.6%	90.9%
P20	Untransformed	Cloglog	68.2%	81.8%	88.6%
P20	Untransformed	Logit	68.2%	77.3%	88.6%
P20	Untransformed	Probit	68.2%	77.3%	88.6%

The table shows the percentage of cases in which the null hypothesis is rejected in the LR test for the state of turbulence models with different assumptions about the distribution of a random error, a set of independent variables and the dependent variable definition. The table shows results of the test for 1%, 5% and 10% significance levels.

Table 4. Results for the Gini coefficient analysis for the individual assets analysis

DEP. VAR.	MODEL	DATA TYPE	Gini
P1	Cloglog	Untransformed	0.853
P1	Logit	Untransformed	0.859
P1	Probit	Untransformed	0.877
P5	Cloglog	Untransformed	0.521
P5	Logit	Untransformed	0.526
P5	Probit	Untransformed	0.536

Table 4. Continued

DEP. VAR.	MODEL	DATA TYPE	Gini
P10	Cloglog	Untransformed	0.379
P10	Logit	Untransformed	0.384
P10	Probit	Untransformed	0.388
P20	Cloglog	Untransformed	0.245
P20	Logit	Untransformed	0.248
P20	Probit	Untransformed	0.249

The table shows the average value of the Gini coefficient.

Table 5. Selection of the optimal cut-off point based on the CROC criterion. The P1 dependent variable

DATA TYPE	DEP. VAR.	CUT-OFF POINT	CROC Cloglog	CROC Logit	CROC Probit
Untransformed	P1	0.01	0.203	0.202	0.185
Untransformed	P1	0.02	0.245	0.236	0.205
Untransformed	P1	0.03	0.307	0.294	0.246
Untransformed	P1	0.04	0.326	0.320	0.293
Untransformed	P1	0.05	0.369	0.363	0.345
Untransformed	P1	0.06	0.404	0.388	0.377
Untransformed	P1	0.07	0.417	0.409	0.413
Untransformed	P1	0.08	0.438	0.426	0.434
Untransformed	P1	0.09	0.469	0.449	0.455
Untransformed	P1	0.1	0.490	0.476	0.497

The table shows the selection of the optimal cut-off point based on the CROC criterion for the models with the P1 dependent variable.

Table 6. Selection of the optimal cut-off point based on the CROC criterion. The P5 dependent variable

DATA TYPE	DEP. VAR.	CUT-OFF POINT	CROC Cloglog	CROC Logit	CROC Probit
Untransformed	P5	0.01	0.817	0.809	0.782
Untransformed	P5	0.02	0.639	0.634	0.622
Untransformed	P5	0.03	0.510	0.506	0.503
Untransformed	P5	0.04	0.438	0.433	0.432
Untransformed	P5	0.05	0.418	0.414	0.412
Untransformed	P5	0.06	0.436	0.432	0.424
Untransformed	P5	0.07	0.468	0.462	0.449
Untransformed	P5	0.08	0.501	0.493	0.486
Untransformed	P5	0.09	0.537	0.534	0.530

The table shows the selection of the optimal cut-off point based on the CROC criterion for the models with the P5 dependent variable.

Table 7. Selection of the optimal cut-off point based on the CROC criterion. The P10 dependent variable

DATA TYPE	DEP. VAR.	CUT-OFF POINT	CROC Cloglog	CROC Logit	CROC Probit
Untransformed	P10	0.06	0.637	0.633	0.631
Untransformed	P10	0.07	0.571	0.569	0.569
Untransformed	P10	0.08	0.523	0.521	0.522
Untransformed	P10	0.09	0.496	0.493	0.493
Untransformed	P10	**0.1**	0.487	0.482	0.482
Untransformed	P10	0.11	0.500	0.496	0.492
Untransformed	P10	0.12	0.524	0.519	0.515
Untransformed	P10	0.13	0.547	0.544	0.540
Untransformed	P10	0.14	0.577	0.573	0.567
Untransformed	P10	0.15	0.609	0.599	0.599

The table shows the selection of the optimal cut-off point based on the CROC criterion for the models with the P10 dependent variable.

Table 8. The best combinations of the state of turbulence models assumptions. Individual assets analysis

DEPENDENT VARIABLE	CUT-OFF POINT	MODEL	DATA TYPE
P10	10%	Logit	Untransformed
P10	10%	Probit	Untransformed
P10	10%	Cloglog	Untransformed
P5	5%	Logit	Untransformed
P5	5%	Probit	Untransformed
P5	5%	Cloglog	Untransformed
P1	1%	Logit	Untransformed
P1	1%	Probit	Untransformed
P1	1%	Cloglog	Untransformed

The table shows the nine combinations of assumptions, which (based on the results obtained) should define high-quality state of turbulence models.

Table 9. Companies, which shares were included in the study for portfolios

NO	PORTFOLIO 1	PORTFOLIO 2	PORTFOLIO 3	PORTFOLIO 4	PORTFOLIO 5
1	ATLANTIS S.A.	AMPLI S.A.	ATLANTA S.A.	ASSECO POLAND S.A.	ATM GROUP S.A.
2	BBI ZENERIS NFI S.A.	FORTE S.A.	AWBUD S.A.	BIOTON S.A.	ZO BYTOM S.A.
3	BIOTON S.A.	INTER GROCLIN AUTO S.A.	DUDA S.A.	ELZAB S.A.	CEZ A.S.
4	ECHO INVESTMENT S.A.	HYDROTOR S.A.	EUROCASH S.A.	GLOBE TRADE CENTRE S.A.	IMPEXMETAL S.A.

Table 9. Continued

NO	PORTFOLIO 1	PORTFOLIO 2	PORTFOLIO 3	PORTFOLIO 4	PORTFOLIO 5
5	ENERGOPOL-POŁUDNIE S.A.	KGHM S.A.	DM IDM S.A.	IMPEXMETAL S.A.	IZOLACJA JAROCIN S.A.
6	GLOBE TRADE CENTRE S.A.	MOSTOSTAL PŁOCK S.A.	PEPEES S.A.	LPP S.A.	NORDEA BP S.A.
7	PKO BP S.A.	POLICE S.A.	PRÓCHNIK S.A.	NOVITA S.A.	ORCO PROPERTY GROUP S.A.
8	TRAVELPLANET.PL S.A.	PULAWY S.A.	REDAN S.A.	PBS FINANSE S.A.	TRAVELPLANET.PL S.A.
9	WISTIL S.A.	SYGNITY S.A.	STOMIL SANOK S.A.	POLCOLORIT S.A.	ULMA S.A.
10	ŻYWIEC S.A.	ZETKAMA S.A.	SUWARY S.A.	SOPHARMA AD	ZELMER S.A.

Table 10. The Hosmer-Lemeshow test results for the portfolio analysis

DEP. VAR.	MODEL	PORT. 1	PORT. 2	PORT. 3	PORT. 4	PORT. 5	AVG
P10	Cloglog	100%	99%	99%	99%	100%	100%
P10	Probit	92%	100%	100%	93%	98%	97%
P10	Logit	98%	100%	100%	99%	100%	99%
P5	Cloglog	88%	100%	100%	100%	97%	97%
P5	Probit	88%	97%	97%	100%	100%	96%
P5	Logit	76%	100%	100%	100%	100%	95%
P1	Cloglog	100%	84%	84%	100%	100%	93%
P1	Probit	100%	100%	100%	100%	100%	100%
P1	Logit	100%	99%	99%	100%	100%	100%

The table shows the percentage of the cases in which there is no basis to reject the null hypothesis in the Hosmer-Lemeshow test for the state of turbulence models with different assumptions about the distribution of a random error, a set of independent variables and the dependent variable definition.

Table 11. The LR test results for the portfolio analysis

DEP. VAR.	MODEL	PORT. 1	PORT. 2	PORT. 3	PORT. 4	PORT. 5	AVG
P10	Cloglog	100%	100%	100%	100%	100%	100%
P10	Probit	100%	100%	100%	100%	100%	100%
P10	Logit	100%	100%	100%	100%	100%	100%
P5	Cloglog	100%	100%	100%	100%	100%	100%
P5	Probit	100%	100%	100%	100%	100%	100%
P5	Logit	100%	100%	100%	100%	100%	100%

Table 11. Continued

DEP. VAR.	MODEL	PORT. 1	PORT. 2	PORT. 3	PORT. 4	PORT. 5	AVG
P1	Cloglog	100%	100%	100%	100%	100%	100%
P1	Probit	100%	100%	100%	100%	100%	100%
P1	Logit	100%	100%	100%	100%	100%	100%

The table shows the percentage of the cases in which the null hypothesis is reject in the LR test for the state of turbulence models with different assumptions about the distribution of a random error, a set of independent variables and the dependent variable definition.

Table 12. Results for the Gini coefficient analysis for the portfolios. In-sample analysis

DEP. VAR.	MODEL	PORT. 1	PORT. 2	PORT. 3	PORT. 4	PORT. 5	AVG
P10	Cloglog	0.43	0.48	0.48	0.46	0.41	0.45
P10	Probit	0.44	0.50	0.50	0.47	0.42	0.47
P10	Logit	0.43	0.49	0.49	0.47	0.41	0.46
P5	Cloglog	0.54	0.59	0.59	0.57	0.59	0.58
P5	Probit	0.56	0.61	0.61	0.59	0.61	0.60
P5	Logit	0.55	0.60	0.60	0.58	0.60	0.59
P1	Cloglog	0.88	0.85	0.85	0.90	0.90	0.88
P1	Probit	0.90	0.87	0.87	0.92	0.91	0.89
P1	Logit	0.89	0.86	0.86	0.91	0.90	0.88

The table shows the average value of the Gini coefficient for the in-sample analysis.

Table 13. Results for the Gini coefficient analysis for the portfolios. Out-of-sample analysis

DEP. VAR.	MODEL	PORT. 1	PORT. 2	PORT. 3	PORT. 4	PORT. 5	AVG
P10	Cloglog	0.39	0.34	0.34	0.38	0.33	0.36
P10	Probit	0.41	0.41	0.41	0.41	0.33	0.40
P10	Logit	0.39	0.37	0.37	0.40	0.33	0.37
P5	Cloglog	0.32	0.45	0.45	0.51	0.46	0.44
P5	Probit	0.36	0.46	0.46	0.57	0.52	0.48
P5	Logit	0.34	0.46	0.46	0.54	0.47	0.45
P1	Cloglog	- 0.89	0.44	0.44	0.30	0.40	0.14
P1	Probit	- 0.57	0.49	0.49	0.31	0.43	0.23
P1	Logit	- 0.86	0.46	0.46	0.29	0.41	0.15

The table shows the average value of the Gini coefficient for out-of-sample analysis.

Table 14. Results for the CROC criterion analysis for the portfolios. In-sample analysis

DEP. VAR.	MODEL	PORT. 1	PORT. 2	PORT. 3	PORT. 4	PORT. 5	AVG
P10	Cloglog	0.47	0.41	0.41	0.43	0.48	0.44
P10	Probit	0.47	0.41	0.41	0.43	0.46	0.44
P10	Logit	0.47	0.41	0.41	0.43	0.47	0.44
P5	Cloglog	0.42	0.35	0.35	0.41	0.38	0.38
P5	Probit	0.41	0.34	0.34	0.39	0.36	0.37
P5	Logit	0.41	0.35	0.35	0.41	0.35	0.38
P1	Cloglog	0.16	0.22	0.22	0.15	0.16	0.18
P1	Probit	0.16	0.21	0.21	0.15	0.15	0.18
P1	Logit	0.16	0.21	0.21	0.15	0.16	0.18

The table shows the selection of the optimal cut-off point based on the CROC criterion for the in-sample analysis.

Table 15. Results for the CROC criterion analysis for the portfolios. Out-of-sample analysis

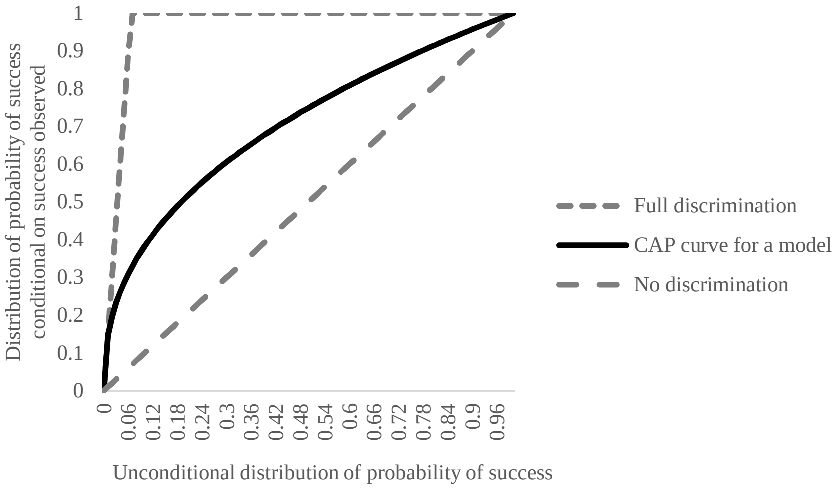
DEP. VAR.	MODEL	PORT. 1	PORT. 2	PORT. 3	PORT. 4	PORT. 5	AVG
P10	Cloglog	0.61	0.64	0.64	0.82	0.61	0.66
P10	Probit	0.64	0.58	0.58	0.71	0.59	0.62
P10	Logit	0.61	0.61	0.61	0.82	0.61	0.65
P5	Cloglog	0.76	0.49	0.49	0.73	0.78	0.65
P5	Probit	0.77	0.49	0.49	0.55	0.78	0.62
P5	Logit	0.76	0.49	0.49	0.64	0.78	0.63
P1	Cloglog	1	1	1	0.67	0.51	0.84
P1	Probit	1	1	1	0.67	0.5	0.84
P1	Logit	1	1	1	0.67	0.51	0.84

The table shows the selection of the optimal cut-off point based on the CROC criterion for out-of-sample analysis.

Table 16. Six the best combinations of the state of turbulence models assumptions

ASSUMPTIONS	DEPENDENT VARIABLE	CUT-OFF POINT	DATA TYPE	MODEL
GROUP 1	P10	10%	Untransformed	Probit
GROUP 2	P10	10%	Untransformed	Logit
GROUP 3	P10	10%	Untransformed	Cloglog
GROUP 4	P5	5%	Untransformed	Probit
GROUP 5	P5	5%	Untransformed	Logit
GROUP 6	P5	5%	Untransformed	Cloglog

Figure 1. Example of CAP curve



Source: own calculation based on: BCBS (2005).