Is it so bad that we cannot recognize black swans?

Fuad Aleskerov, Lyudmila Egorova

National Research University – Higher School of Economics

Analyzing the reasons of financial crises in «The Black Swan: The Impact of The Highly Improbable» Nassim Nicolas Taleb concludes that modern economic models badly describe reality for they are not able to forecast such crises in advance.

We tried to present processes on stock exchange as two random processes one of which happens rather often (regular regime) and the other one – rather rare. Our answer is that if regular processes are correctly recognized with the probability higher than $\frac{1}{2}$, this allows to get positive average gain. We believe that this very phenomenon lies in the basis of unwillingness of people to expect crises permanently and to try recognizing them.

Keywords: Black Swans, financial crises, Poisson process

1. Introduction

Analyzing the reasons of financial crises in «The Black Swan: The Impact of The Highly Improbable» [6] Nassim Nicolas Taleb concludes that modern economic models badly describe reality for they are not able to forecast such crisis in advance. All extraordinary events, e.g. crises, are named by the author “The Black Swans”. Let us give some quotes.

«Before the discovery of Australia, people in the old world were convinced that all swans were white, an unassailable belief as it seemed completely confirmed by empirical evidence. The sighting of the first black swan might have been an interesting surprise for a few ornithologists (and others extremely concerned with the coloring of birds), but that is not where the significance of the story lies. It illustrates a severe limitation to our learning from observations or experience and the fragility of our knowledge. One single observation can invalidate a general statement derived from millennia of confirmatory sightings of millions of white swans. All you need is one single (and, I am told, quite ugly) black bird.

… What we call here a Black Swan (and capitalize it) is an event with the following three attributes. First, it is an outlier, as it lies outside the realm of regular expectations, because nothing in the past can convincingly point to its possibility. Second, it carries an extreme impact. Third, in spite of its outlier status,
human nature makes us concoct explanations for its occurrence after the fact, making it explainable and predictable.

I stop and summarize the triplet: rarity, extreme impact, and retrospective (though not prospective) predictability. A small number of Black Swans explain almost everything in our world, from the success of ideas and religions, to the dynamics of historical events, to elements of our own personal lives. Ever since we left the Pleistocene, some ten millennia ago, the effect of these Black Swans has been increasing. It started accelerating during the industrial revolution, as the world started getting more complicated, while ordinary events, the ones we study and discuss and try to predict from reading the newspapers, have become increasingly inconsequential.

Just imagine how little your understanding of the world on the eve of the events of 1914 would have helped you guess what was to happen next. (Don’t cheat by using the explanations drilled into your cranium by your dull high school teacher). How about the rise of Hitler and the subsequent war? How about the precipitous demise of the Soviet bloc? How about the rise of Islamic fundamentalism? How about the spread of the Internet? How about the market crash of 1987 (and the more unexpected recovery)? Fads, epidemics, fashion, ideas, the emergence of art genres and schools. All follow these Black Swan dynamics. Literally, just about everything of significance around you might qualify.

This combination of low predictability and large impact makes the Black Swan a great puzzle; but that is not yet the core concern of this book. Add to this phenomenon the fact that we tend to act as if it does not exist! I don’t mean just you, your cousin Joey, and me, but almost all “social scientists” who, for over a century, have operated under the false belief that their tools could measure uncertainty. For the applications of the sciences of uncertainty to real-world problems has had ridiculous effects; I have been privileged to see it in finance and economics. Go ask your portfolio manager for his definition of “risk,” and odds are that he will supply you with a measure that excludes the possibility of the Black
Swan—hence one that has no better predictive value for assessing the total risks than astrology (we will see how they dress up the intellectual fraud with mathematics). This problem is endemic in social matters.

The central idea of this book concerns our blindness with respect to randomness, particularly the large deviations: Why do we, scientists or nonscientists, hotshots or regular Joes, tend to see the pennies instead of the dollars? Why do we keep focusing on the minutiae, not the possible significant large events, in spite of the obvious evidence of their huge influence?» (Prologue, pp. XVII-XIX).

Thus, the author confirms that modern science almost doesn’t have tools to predict such unusual events. Moreover, he gives the following rather sharp characterization of modern economic knowledge.

«In orthodox economics, rationality became a straitjacket. Platonified economists ignored the fact that people might prefer to do something other than maximize their economic interests. This led to mathematical techniques such as “maximization,” or “optimization,” on which Paul Samuelson built much of his work… It involves complicated mathematics and thus raises a barrier to entry by non-mathematically trained scholars. I would not be the first to say that this optimization set back social science by reducing it from the intellectual and reflective discipline that it was becoming to an attempt at an “exact science.”» (p.184).

And, finally,

«… those who started the game of “formal thinking,” by manufacturing phony premises in order to generate “rigorous” theories, were Paul Samuelson, Merton’s tutor, and, in the United Kingdom, John Hicks. These two wrecked the ideas of John Maynard Keynes, which they tried to formalize (Keynes was interested in uncertainty, and complained about the mind-closing certainties induced by models). Other participants in the formal thinking venture were Kenneth Arrow and Gerard Debreu. All four were Nobeled… All of them can be
safely accused of having invented an imaginary world, one that lent itself to their mathematics.» (p.283).

Not being adherents of any particular points of view\textsuperscript{1} in the economic community, we have tried to present the processes occurring on the exchange in the form of two random processes, one of which occurs frequently (normal mode) and the other – rarely (crisis).

Next, we estimated the average gain with the different probabilities of correct recognition of these processes and used the resulting estimates for the actual processes on the exchange.

Briefly, we’ve got the following answer: if frequent, regular processes are detected correctly even with a probability higher than $\frac{1}{2}$, it almost always allows to have a positive average gain. This very phenomenon seems underlies the reluctance of people to expect crises all the time and do not try to identify them.

Also we extended basic mathematical model allowing player to learn on his/her own behavior and to receive an award for the ‘correct’ behavior. Each of the proposed new models allows the player to have more freedom in his/her decisions and make mistakes in rare events more often.

The structure of the paper is as follows. Section 2 proposes the statement of the problem and its solution, Sections 3 investigates the constraints for nonnegative gain. Section 4 deals with application of the model to real data, Section 5 proposes the new models, and Section 6 concludes. In Appendix 1 we derive the formula and proofs of the theorems; in Appendix 2 we consider the estimates of parameters for the various stock market indices.

\textsuperscript{1} Although we must say that one of the authors of this work – F. Aleskerov – met Professor Paul Samuelson in 1984, and Professor Kenneth Arrow has always had high influence to him. Meetings with Professor Arrow are always the great intellectual feast.
Acknowledgements. F.Aleskerov expresses sincere gratitude to Professor Andrey Yakovlev, who presented him the book [6]; disagreement with some of the conclusions of this book has stimulated the creation of the proposed models. The authors are also grateful to Dmitriy Golembiovsky, Elena Goryainova, Alexander Lepskii, Peter Hammond, Kanak Patel, Henrik Penikas, Vladislav Podinovskii, Christian Seidl for valuable remarks and comments.

We are grateful also for partial financial support of the Laboratory DECAN of NRU HSE and NRU HSE Science Foundation (grant № 10-04-0030).

F. Aleskerov thanks also Magdalene College of University of Cambridge during the stay in which this work was completed. He appreciates the hospitality and kindness of Dr. Patel, the Fellow of Magdalene College.

2. Problem

The flow of events of two types – type \( Q \) (from \( \text{quick} \)) and type \( R \) (from \( \text{rare} \)) – enters the device. Each of them is the simplest, i.e. stationary, ordinary and has no aftereffects [2]. The intensity of the flow of events of type \( Q \) is equal to \( \lambda \), the intensity of the flow of events of \( R \) is equal to \( \mu \), where \( \lambda \gg \mu \) (\( Q \)-type events are far more frequent than the \( R \)-type events).

The problem of the device is to recognize coming event \( X \). If an event \( Q \) occurs and device identified it correctly, then it would promote getting a small reward \( a \), if the error occurred, and the event \( Q \) has been recognized as the event \( R \), then the device is ‘fined’ by an amount \( b \). The probabilities of such outcomes are known and equal \( p_1 \) and \( q_1 \), respectively. Similarly for the events of the type \( R \) – correct identification of coming event \( R \) will give the value of \( c \), where \( c \gg a \), and incorrect recognizing will give loss – \( d \), \( d \gg b \). After each coming event received values of ‘win’ / ‘loss’ are added to the previous amount (Fig. 1).

How large on average will be the amount received for the time \( t \)?
One possible implementation of a random process $Z(t)$, which is equal to the sum of all values of a random variable $X$ received at the time $t$, is given on Fig. 2.

Random value $Z$ of the total sum of the received prizes during the time $t$ is a compound Poisson type variable since the number of terms in the sum $Z = \sum X_i$ is also a random variable and depends on the flow of events received by the device.

**Theorem 1.** The expected value of $Z$ is equal to

$$E[Z] = (\lambda(p_1a - q_1b) + \mu(p_2c - q_2d))t.$$

The proofs of this and next statements are given in Appendix 1.
3. Analysis of the solution

Let us require the expectation of $Z$ to be non-negative. The values $q_1$, $q_2$ are known and predetermined. Then let us find for which values of other parameters the requirement holds

$$E(Z) = \left((1 - q_1)\lambda a - q_1\lambda b + (1 - q_2)\mu c - q_2\mu d\right) \cdot t \geq 0$$

As $t > 0$ and $\mu > 0$, then $E(Z) \geq 0$ holds when

$$(a - q_1 a - q_1 b)\frac{\lambda}{\mu} + (c - q_2 c - q_2 d) \geq 0.$$  

Denoting $E(W) = a - q_1 a - q_1 b$ is an expected gain of only $Q$-event and $E(Y) = c - q_2 c - q_2 d$ is an expected gain from only $R$-type event, we can set the Table 1, where you can see how the solution $E(Z) \geq 0$ depends on these parameters.

**Table 1. The conditions of non-negativity of the total gain**

<table>
<thead>
<tr>
<th>$E(Y)$</th>
<th>$E(W)$</th>
<th>$E(Z)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0</td>
<td>&lt; 0</td>
<td>for $E(Z) \geq 0$ required $\frac{\lambda}{\mu} \geq -\frac{E(Y)}{E(W)}$</td>
</tr>
<tr>
<td>= 0</td>
<td>$E(Z) &lt; 0$</td>
<td>$E(Z) &lt; 0$</td>
</tr>
<tr>
<td>&gt; 0</td>
<td>for $E(Z) \geq 0$ required $\frac{\lambda}{\mu} \leq -\frac{E(Y)}{E(W)}$</td>
<td>$E(Z) \geq 0$</td>
</tr>
</tbody>
</table>

Let $\lambda = 246, \mu = 4$. If $E(W)$ – the expected value of gain from the regular event $Q$ – is assumed to be equal to 1, then what should be the expected value $E(Y)$ of the $R$-type event?

According to Table 1, the following inequality $E(Y) \geq -61.5$ holds, i.e., if we have very small expected gain in one unit under frequent observations, we can cover very large losses which happen quite rare.
Now consider the case when the parameters \( q_1 \) and \( q_2 \) are unknown. What conditions should satisfy the values of \( q_1 \) and \( q_2 \) for the expected value \( E(Z) \) to be nonnegative with all other parameters being fixed?

Both \( q_1 \) and \( q_2 \) are the probabilities of incorrectly recognized events \( Q \) and \( R \), so we have to solve the system of inequalities

\[
\begin{aligned}
E(Z) &= [\lambda((1 - q_1)a - q_1b) + \mu((1 - q_2)c - q_2d)]t \geq 0, \\
0 &\leq q_1 \leq 1, \\
0 &\leq q_2 \leq 1,
\end{aligned}
\]

when the restrictions on the parameters are \( a, b, c, d \geq 0, \lambda, \mu, t > 0, a + b > 0, c + d > 0 \) (the latter two inequalities mean that both \( a \) and \( b, c \) and \( d \) cannot be equal to zero since the cases with the events with zero losses and gains simultaneously are not interesting)

\[
\begin{aligned}
q_1\lambda(a + b) + q_2\mu(c + d) &\leq \lambda a + \mu c, \\
0 &\leq q_1 \leq 1, \\
0 &\leq q_2 \leq 1.
\end{aligned}
\]

The solution is the range of values \( q_1 \) and \( q_2 \) defined via the system of inequalities

\[
\begin{aligned}
0 &\leq q_1 \leq \min\left\{1, \frac{\lambda a + \mu c}{\lambda(a + b)}\right\}, \\
0 &\leq q_2 \leq 1 \text{ if } 0 \leq q_1 \leq \min\left\{1, \frac{\lambda a - \mu d}{\lambda(a + b)}\right\}, \\
0 &\leq q_2 \leq -\frac{\lambda(a + b)}{\mu(c + d)}q_1 + \frac{\lambda a + \mu c}{\mu(c + d)} \text{ if } \max\left\{0, \frac{\lambda a - \mu d}{\lambda(a + b)}\right\} \leq q_1 \leq \min\left\{1, \frac{\lambda a + \mu c}{\lambda(a + b)}\right\}.
\end{aligned}
\]

Depending on the values of \( q_1 \) and \( q_2 \) the corresponding areas look as follows (Fig. 3)
Let $q_2$ be equal to 1 (it means that the player can’t recognize crises at all). How often the player can fail to detect regular event to have still positive or at least zero average gain $E(Z)$?

If we take intensities equal to $\lambda = 246, \mu = 4$, and the single wins as $a = 0.6, -b = -0.6, c = 2.8, -d = -2.9$ the answer will be $q_1 = 0.46$.

We call it a critical value of error probability in ordinary events $Q$.

4. Application to real data

Let the device be a stock exchange and events $Q$ and $R$ describe a ‘quiet life’ and a ‘crisis’, respectively. According to the model, events $Q$ occur more frequently than $R$ that corresponds to the fact that the crises in our lives are fortunately rare.

The event $X$ can be interpreted as a signal received by a broker about the changes of the economics that helps him to decide whether the economy is in ‘a normal mode’ or in a crisis. For example, does the fall of oil prices mean the

![Fig. 3. Examples of possible areas of parameters](image-url)
beginning of the recession in economy or it is a temporary phenomenon and will not change the economy at all.

The values \(a, b, c, d\) also have some meaning in such interpretation. If the event \(Q\) occurs (which means that the economics is stable), and broker correctly recognizes it, then he can get a small income (value \(a\)). If the event \(Q\) will be taken instead of \(R\), he will lose the amount of \(-b\). If the \(R\)-event occurred (crisis) and it was not recognized correctly, the broker will lose more (value \(-d\)). If he could forecast a crisis, he can earn a good deal of money on this - correct identification of the event of type \(R\) gives the broker the value \(c\).

Such outcomes correspond to the opening of the long and short positions in a period of growth and recession in the work of the trader. A long position means that the broker buys assets to sale some time later at a higher price. A short position means that the broker sells assets with the hope of further buying at a lower price.

A long position will bring a small income \(a\) and a significant loss of \(-d\) to the trader, when the market is growing (‘regular’ event) and falls (‘crisis’), respectively. It will be the opposite with the short positions: trader will lose some value \(-b\) in case of economic growth, but he can earn a considerable amount of \(c\) in the case of strong fall in the crisis.

The basic model contains some important assumptions about the kind of processes. It is known that the actual flows of stock exchange events are not simplest. It can be better described as a piecewise stationary stochastic processes with unknown switch points. Stationarity of real data for S&P500 has been tested and periods of crisis and periods of absence of shocks show stationary time series indeed (we used time series of returns of the stock index as in small samples about 10-20 points and for a long period of several years corresponding to the only regular or the only crises days; the Dickey-Fuller Unit root test and analysis of autocorrelation and partial autocorrelation functions were used to control of stationary property). However, we cannot speak of stationary series in the long time. Aware of these shortcomings of our model, we consider it as a first step to
the study of real stock exchange and the basis for constructing more sophisticated models. Another assumption is that we do not imply that \( R \)-type events reflects only crises as recession of economy. According to Taleb’s book the black swans are unpredictable but not necessary ‘bad’ events.

For estimation of the parameters of the model we will use time series of the stock index S&P 500. S&P500 includes 500 selected stocks of the USA having the largest capitalization. The list is owned and formed by Standard & Poor's [7].

The time interval has been taken for over 10 years - from August 1999 until December 2009. We took the mean of the value of opening and closing as the index value for the day (Fig. 4). The same analysis was done with the closing prices (see Appendix 2).

![Fig. 4. Stock index S&P500. Some points are marked with their date on the left.](image)

What value can be an indicator of market behavior and a signal of crisis?

We use the volatility of the index because the fall of the index immediately reflects on it (on its amplitude). We evaluated the volatility of the index with a sliding interval of 20 days to find when the problems occur in the economy and
the recession begins, and took the previous 20 index values $S_i$ for each day and calculate the standard deviation of the sample mean $\tilde{S}$

$$\sigma_j = \sqrt{\frac{1}{19} \sum_{i=j-20}^{j} (S_i - \tilde{S})^2}, j = 21, 22, \ldots$$

Then the values of volatility were divided to the value of the index to make these quantities dimensionless. We accept that if the volatility is greater than the threshold, this means the occurrence of an event of $R$-type, i.e. a crisis. Figure 5 shows that high values of volatility amplitude correspond to the drop of the index. We choose 6% as a threshold.

The alternative method to estimate the number of $R$-type crises using the index return you can see in Appendix 2.

Fig. 5. Stock index S&P500. The solid line corresponds to the value of S&P500 and dotted line corresponds to its volatility. The highest peaks of volatility marked with their date on the left.

And we can evaluate the other model parameters using the return data $r_i = \frac{S_i - S_{i-1}}{S_{i-1}} \cdot 100\%, i = 2, \ldots, 2665$. We estimated $a$ and $b$ at points corresponding
to the event \( Q \) (using the value of the index volatility – it should be less than the threshold value for the \( Q \)-event). If the index goes up at this moment (the volatility is positive and less than the threshold), it means that the event \( a \) was realized, and if it goes down – then \( -b \) is observed. The same approach was used for \( R \)-events, i.e. if the index has increased in compare with the previous value, the change of the index is \( c \); if the index has fallen, the change is \( -d \).

Then the model parameters will be \( \lambda = 246, \mu = 4, a = 0.6, -b = -0.6, c = 2.8, -d = -2.9 \). Since we took daily prices, these intensities reflect the fact that taking 250 working days in a year we have 4 crises days (when the daily volatility is higher than 6%).

One can see which should be the probability of error in the identification process for the expected gain of broker to be non-negative under these values of parameters. For the 4% threshold the corresponding region in the \( q_1 - q_2 \) plane is shown on Fig. 7.

![Fig. 7. Dashed area shows when the expected gain \( E(Z) \) is nonnegative](image)

As it is seen, it is enough to identify \( Q \)-events better than in half of the cases to ensure a positive outcome. Indeed, if we choose the horizon of 1 year and the error probability for events \( Q \) and \( R \) being \( q_1=0.46, q_2=1 \) (i.e. even when crises are not recognized correctly at all), the expected gain is still positive. Naturally, with \( q_1 \) and \( q_2 \) decreasing, the average gain increases.

If the player do not try to recognize coming events but ‘toss up a coin’ to decide (this corresponds to the model with \( q_1 = \frac{1}{2}, q_2 = \frac{1}{2} \)), then he/she will have negative gain \( E(Z) = -0.2\% \) in a year.
The same calculations can be done for other indices (Table 2).

Table 2. Parameters for other indices with the threshold 6%

<table>
<thead>
<tr>
<th>Parameters, %</th>
<th>$a$</th>
<th>$-b$</th>
<th>$c$</th>
<th>$-d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500</td>
<td>0.6</td>
<td>-0.6</td>
<td>2.8</td>
<td>-2.9</td>
</tr>
<tr>
<td>Dow Jones</td>
<td>0.6</td>
<td>-0.6</td>
<td>1.9</td>
<td>-2.5</td>
</tr>
<tr>
<td>CAC 40</td>
<td>0.8</td>
<td>-0.8</td>
<td>3.0</td>
<td>-2.5</td>
</tr>
<tr>
<td>DAX</td>
<td>0.8</td>
<td>-0.9</td>
<td>2.1</td>
<td>-2.5</td>
</tr>
<tr>
<td>Nikkei 225</td>
<td>0.8</td>
<td>-0.9</td>
<td>2.6</td>
<td>-3.2</td>
</tr>
<tr>
<td>Hang Seng</td>
<td>0.9</td>
<td>-0.9</td>
<td>2.6</td>
<td>-3.0</td>
</tr>
</tbody>
</table>

We can see that the relative parameter estimates for the indices are approximately equal – index varies with the scope of 0.8-0.9% in the quiet period and its amplitude increases approximately 4 times in time of crisis.

It should be noted that these data should reflect the change in the index in percentages compared to the previous day, rather than racing the index for one day. Daily changes may be more significant, for example, in April 17, 2000 the difference between the highest and the lowest index value of Japanese Nikkei 225 amounted to almost 100 points (approximately 10%). However, April 16 drop was only 3.8% in comparison with the Nikkei the previous day.

5. New models

Let us extend the basic model adding new conditions.

Since the intensity of regular events $Q$ is much higher than the intensity of rare events, regular events often happen one by one and form a sequence of these ‘peaceful’ events. So, we can suggest that the device can ‘learn’ on such sequences and turn them to its advantage raising the winnings from regular events.

It means that if the event $Q$ has been detected correctly by the device consecutively $k$ times, then it gets a higher award $a + \varepsilon$ (not $a$ as in the basic model) for the recognition of the $Q$-events.
The experience function $S_i$ on step $i$ is a random variable equal to the number of consequently correctly recognized $Q$-events (we designate an event of correct recognition of $Q$-type event as $A$) that occurred by this step

$$S_0 = 0,$$

$$S_i = \begin{cases} S_{i-1} + 1, & \text{if } A \text{ occurs}, \\ 0, & \text{if } \overline{A} \text{ occurs (}A\text{ didn’t occur).} \end{cases}$$

The model is graphically depicted on Fig.8.

![Diagram](image)

**Fig. 8.** The graph of model with stimulation

The experience function $S_i(t)$ for implementation of random process $Z(t)$ from Fig.2 is shown on Fig. 9.
Fig. 9. One possible implementation of $Z(t)$ and its experience function $S_i(t)$

**Theorem 2.** The expectation of total gain in the model with stimulation is

$$E(Z) = E(X^{(1)}) \cdot \left[ \lambda_{NZ} \cdot \frac{\Gamma(k-1, \lambda_{NZ})}{\Gamma(k-1)} + (k-1) \cdot \left( 1 - \frac{\Gamma(k, \lambda_{NZ})}{\Gamma(k)} \right) \right] +$$

$$+ E(X^{(2)}) \cdot \left[ \lambda_{NZ} \cdot \left( 1 - \frac{\Gamma(k-1, \lambda_{NZ})}{\Gamma(k-1)} \right) + (1-k) \cdot \left( 1 - \frac{\Gamma(k, \lambda_{NZ})}{\Gamma(k)} \right) \right], \quad (5.1)$$

where $\lambda_{NZ} = (\lambda + \mu)t$ is the intensity of flow of unknown events (both $Q$ and $R$), $X_i^{(1)}$ and $X_i^{(2)}$ are the random variables for the total gain in case $i < k$ and $i \geq k$, respectively. Their distributions can be obtained by selection of relevant cases from the distribution law of the general variable $X_i$.
Let us consider more complicated model where our device will also learn on his actions: if $Q$-event was successfully recognized $k$ times consequently (it means that $k$ times the device received an award $a$), then it will further recognize an event of $Q$ correctly with greater probability $p_1^* = p_1 + \delta > p_1$.

Denote $S_i$ as an experience function on step $i$, it is a random variable of the number consequently correctly recognized $Q$-events. This experience function is defined almost like an experience function in the previous model, but this function is changed if an event $X_i = a$ occurs, i.e. our device successfully detected coming event $Q$

$$S_0 = 0,$$

$$S_i = \begin{cases} S_{i-1} + 1, & \text{if } a \text{ occured,} \\ 0, & \text{if } \bar{a} \text{ occured (}a\text{ has not occured).} \end{cases}$$

The graph for such model is given on Fig. 10.
The question is still about the expected value of the total gain, but now we have to know the probabilities \( P\{S_i < k\} \) and \( P\{S_i \geq k\} \), because now the random variable of single winning \( X_i \) takes values \(-d, -b, a, c\) with probabilities

\[
\Pr\{X_i = x\} = \begin{cases} 
  p_R q_2, & \text{if } x = -d, \\
  p_Q q_1 \Pr\{S_i < k\} + q_i^* \Pr\{S_i \geq k\}, & \text{if } x = -b, \\
  p_Q p_1 \Pr\{S_i < k\} + p_1^* \Pr\{S_i \geq k\}, & \text{if } x = a, \\
  p_R p_2, & \text{if } x = c.
\end{cases}
\]

**Theorem 3.** The probability \( P\{S_i < k\} \) is equal to

\[
P\{S_i < k\} = p_Q (q_1 - q_i^*) \sum_{j=0}^{k-1} P\{S_{i-1-j} < k\} (p_Q p_1)^j + \frac{p_Q q_i^* + p_R}{p_Q q_1 + p_R} (1 - (p_Q p_1)^k).
\]

**Theorem 4.** The sequence \( P\{S_i < k\} \) has the limit

\[
\lim_{i \to \infty} P\{S_i < k\} = \frac{(p_Q q_i^* + p_R)(1 - (p_Q p_1)^k)}{(p_Q q_1 + p_R) - p_Q (q_1 - q_i^*)(1 - (p_Q p_1)^k)}.
\]

Fig. 11 illustrates the sequence of \( P\{S_i < k\} \) and its limit with the example when \( k = 8, q_1 = 0.3, q_1 = 1, \delta = 0.2.\)
Fig. 11. Number of events \( i \) occurred is on the axis OX and probabilities \( P\{S_i < k\} \) are on the axis OY, the line is defined by \( P\{S_i < k\} \) from formula (4.1).

We can use the formula (5.1) to compute the expected gain in this model.

6. Top 10

Applying this model to the stock index S&P 500 gives us a generalized result as the player put his money in equal proportions to all 500 companies from the list of S&P 500. Let us apply the same approach to the companies themselves.

We choose the top 10 companies by capitalization from the S&P 500 list [8]:

1) The Exxon Mobil Corporation, or ExxonMobil, is an American multinational oil and gas corporation;

2) Microsoft Corporation is an American public multinational corporation that develops, manufactures, licenses, and supports a wide range of products and services predominantly related to computing through its various product divisions;


4) JPMorgan Chase & Co. is a global securities, investment banking and retail banking firm;
5) Procter & Gamble Co. is an American multinational corporation that manufactures a wide range of consumer goods;

6) Johnson & Johnson is a global American pharmaceutical, medical devices and consumer packaged goods manufacturer founded in 1886;

7) Apple Inc. is an American multinational corporation that designs and markets consumer electronics, computer software, and personal computers;

8) AT&T Inc. is the largest provider of fixed telephony in the United States, and also provides broadband and subscription television services;

9) International Business Machines (IBM) is a United States multinational technology and consulting firm. IBM manufactures and sells computer hardware and software, and it offers infrastructure, hosting and consulting services in areas ranging from mainframe computers to nanotechnology;

10) Bank of America Corporation is a financial services company, the largest bank holding company in the United States, by assets, and the second largest bank by market capitalization.

We qualify $Q$- and $R$-events using the S&P 500 data, as it was done in Section 4, and then evaluate single winnings $a, b, c, d$ for all 10 firms: if its stock goes up/down when the volatility of S&P 500 is less than 6 % than it is a realization of $a/b$; and if its stock goes up/down when the volatility of S&P 500 is greater than 6 % than it is a realization of $c/d$.

The estimates you can see on the Table 3. The first row shows the parameters of S&P 500 and filled cells show what companies have stocks more volatile than the stock index S&P 500 for $Q$- and $R$- events respectively (if all parameters in absolute value are greater than the same for S&P 500). All of these companies have winnings from $Q$-events greater than the index S&P500, and some of them also have greater parameters of crises events.

Table 3. Top 10 companies (threshold 6%, $\lambda=246$, $\mu=4$)
The analysis with the other thresholds you can find in the Appendix 2.

7. Conclusion

One of our main conclusions is that there is no need to live in the paradigm of an impending crisis. First of all, because it is impossible to predict the time of the crisis. Accordingly, it is impossible to live permanently in the pressure of expectations of the crisis.

In the second place, there is no such need because in a series of regular mass events recognition of such events is much easier and, as we showed, the exact recognition of all process does not really matter.

In the third place, as pointed out by Norbert Wiener, the stock exchange is based on man’s decisions and the prediction of his behavior will lead to the closing of the stock exchange or he will change his behavior strategy, so any attempt to predict are senseless.

Defending economics we can say also that in [5] both crises of 2000 and 2005 were forecast, however, this book did not attract much attention of the scientific community.

In [3] it is pointing out that “…despite its title, Taleb’s book mostly is about how statistical models, especially in finance, should pay more attention to low probability gray swans. It would be much more interesting – though much more challenging – to discuss truly aberrant black swans events to which no probabilities
are attached because the model we use does not even contemplate their possibility.”

Instead of analyzing such probabilities, we showed using very simple model that with a small reward for the correct (with probability higher ½) recognition of the ordinary events (and if crisis events are detected with very low probability) the average player's gain will be positive. In other words, players do not need to play more sophisticated games, trying to identify crises events in advance.

This conclusion resembles the logic of precautionary behavior, that prescripts to play the game with almost reliable small wins.

And we considered new models adding award for ‘successful behavior’ as increase in gain and as increase in the probability of correct recognition, which means that the player can train on his past actions and accumulate experience. Both of these models allows the player to enlarge the total gain and to make more mistakes, because she can get more in the sequence of correctly detected events.
Appendix 1. Proofs of the theorems

Proof of Theorem 1.

The device does not know what event came at a time $i$, so received gain from recognition that event is a random variable $X_i$ with discrete law of distribution. Since the flows of $Q$-events and $R$-events are simplest (and hence stationary), and probabilities $p_1$ and $p_2$ do not depend on time, all $X_i$ is distributed equally as the random variable $X$ with the law of distribution

$$
\Pr\{X = x\} = \begin{cases}
    p_d, & \text{if } x = -d, \\
    p_b, & \text{if } x = -b, \\
    p_a, & \text{if } x = a, \\
    p_c, & \text{if } x = c.
\end{cases}
$$

As flows of $Q$- and $R$-events are simplest, i.e. stationary, ordinary and has no aftereffects, then superposition of these flows will also be a simplest flow with intensity $\lambda + \mu$ [2]. Hence, the probability that coming unknown event is $Q$ is equal to $p_Q = \frac{\lambda}{\lambda + \mu}$ and the probability that coming unknown event is $R$ is equal to $p_R = \frac{\mu}{\lambda + \mu}$. Then $p_d$ (the probability that the random variable $X$ takes the value $-d$) is equal to the probability that the event occurred is the $R$-type and the device has not recognized it, i.e. $p_d = p_R q_2$. We can find other probabilities similarly.

Note that $p_1$ and $p_2$ are independent and they are not probabilities of player’s decisions. The probability that the player will say $Q$ about coming event $X$ is equal to $p_Q p_1 + p_R q_2$ and the opposite probability of choosing $R$ is equal to $p_Q q_1 + p_R p_1$.

Let $F_X(x)$ be the distribution function of a payoff $X$. The total value of received payoffs for time $t$ equals to

$$
Z = \sum_{i=1}^{N_Z} X_i,
$$

where all $X_i$ are random variables of gain of one event and they have the same distribution by the law of distribution of $X$, and $N_Z$ is the number of events.
occurred during the time $t$, it is distributed according to the Poisson distribution with parameter $(\lambda + \mu)t$ (for the flow of events is the simplest flow with the intensity $\lambda + \mu$).

This sum of a Poisson number $N_Z$ terms, where $N_Z$ and $X_i$ are independent, is called a compound Poisson random variable. Its distribution is given by a pair of $P((\lambda + \mu)t; F_Z(x))$, and the explicit form of the distribution function can be obtained by applying the formula of total probability with hypotheses \( \{N_Z = m\} \)

$$F_Z(x) = \sum_{m=0}^{\infty} P\{X_1 + \cdots + X_m \leq x\}P\{N_Z = m\} = \sum_{m=0}^{\infty} F_X^{(m)}(x) \frac{((\lambda + \mu)t)^m}{m!} e^{-\lambda t},$$

where $P\{N_Z = m\} = P_m(t) = \frac{((\lambda+\mu)t)^m}{m!} e^{-(\lambda+\mu)t}$, $F_X^{(m)}(x)$ – $m$-fold convolution of $F_X(x)$, $F_X^{(m)} = F_X^{(m-1)} * F * F$ – the distribution law of variable $X_1 + X_2$ with probabilities $P(X_1 + X_2 = s_j) = \sum_{k=1}^{j} P(X_1 = x_k)P(X_2 = s_j - x_k)$.

Then the expected value of $Z$ is equal to

$$E[Z] = \sum_{j=0}^{\infty} E[Z|N_Z = j]P\{N_Z = j\} = E[X] \sum_{j=0}^{\infty} jP\{N_Z = j\} = E[X] E[N_Z] =$$

$$= \left( -q_2 \frac{d\mu}{\lambda + \mu} - q_1 \frac{b\lambda}{\lambda + \mu} + p_1 \frac{a\lambda}{\lambda + \mu} + p_2 \frac{c\mu}{\lambda + \mu} \right) (\lambda + \mu)t =$$

$$= (\lambda (p_1 a - q_1 b) + \mu (p_2 c - q_2 d))t.$$ 

Proof of Theorem 2.

Suppose $A$ is an event of correct recognition of $Q$-type events and the probability of $A$ is $p_a = p_Q p_1 = \frac{\lambda}{\lambda + \mu} p_1$. If the $Q$-event comes, we should choose the value of winning $a$ or $a + \varepsilon$ according to the experience function $S_i$. We defined the experience function $S_i$ on $i$ step in the following way
\[ S_0 = 0, \]
\[ S_i = \begin{cases} 
  S_{i-1} + 1, & \text{if } A \text{ comes}, \\
  0, & \text{if } A \text{ comes (does not come } A) 
\end{cases} \]

The experience function on step \( i \) can take values from 0 to \( i \) with some probabilities. For example, the probability of \( S_i = i \) is equal \( P\{S_i = i\} = p_a^i \), for another values \( k = 0, 1, 2, \ldots, i - 1 \) the probabilities are \( P\{S_i = k\} = p_a^k \cdot (1 - p_a) \).

Obviously \( p_s = \Pr\{S_i < k\} = 1 \) for \( i < k \). For \( i \geq k \)

\[ p_s = 1 - \Pr\{S_i \geq k\} = 1 - \Pr\{S_i = k\} - \Pr\{S_i = k + 1\} - \cdots - \Pr\{S_i = i\} = \\
= 1 - p_a^k. \]

So, the probability \( p_s \) is equal to

\[ p_s = \begin{cases} 
  1, & \text{if } i < k, \\
  1 - p_a^k, & \text{if } i \geq k. 
\end{cases} \]

Let \( X_i \) be a random variable of gain in the model with stimulation. \( X_i \) depends on number \( i \): for \( i < k \) the probability to get value \( a + \varepsilon \) is zero and for \( i \geq k \) this probability is positive. Then \( X_i \) takes values \(-d, -b, a, a + \varepsilon, c\) with probabilities

\[
\Pr\{X_i = x\} = \begin{cases} 
  p_R q_2, & \text{if } x = -d, \\
  p_Q q_1, & \text{if } x = -b, \\
  p_Q p_1, & \text{if } x = a \text{ and } i < k, \\
  p_Q p_1 \left(1 - (p_Q p_1)^k\right), & \text{if } x = a \text{ and } i \geq k, \\
  0, & \text{if } x = a + \varepsilon \text{ and } i < k, \\
  (p_Q p_1)^{k+1}, & \text{if } x = a + \varepsilon \text{ and } i \geq k, \\
  p_R p_2, & \text{if } x = c. 
\end{cases}
\]

It will be convenient to divide random variable \( X_i \) into two variables \( X_i^{(1)} \) for \( i < k \) and \( X_i^{(2)} \) for \( i \geq k \). Their distributions can be obtained from the law of the random variable \( X_i \) by selection the relevant cases.
Let $\lambda_{NZ} = (\lambda + \mu)t$ be intensity of flow of unknown events. Then for compound Poisson random variable of total winnings $Z = \sum_{i=1}^{N_Z} X_i$ we have

$$E[Z] = \sum_{j=0}^{\infty} E[Z|N_Z = j]P\{N_Z = j\} =$$

$$\sum_{j=0}^{k-1} E\left(\sum_{i=1}^{N_Z} X_i^{(1)} \mid N_Z = j\right)P\{N_Z = j\} + \sum_{j=k}^{\infty} E\left(\sum_{i=1}^{N_Z} X_i^{(1)} \mid N_Z = j\right)P\{N_Z = j\} +$$

$$+ \sum_{j=k}^{\infty} E\left(\sum_{i=k}^{N_Z} X_i^{(2)} \mid N_Z = j\right)P\{N_Z = j\}.$$

The sum is divided into two parts in the last formula because before $k$th term all $X_i$ are $X_i^{(1)}$ and after $k$th term all of them are equal to $X_i^{(2)}$. We take $k > 1$, for $k = 1$ is the case of basic model.

The first sum is

$$\sum_{j=0}^{k-1} E\left(\sum_{i=1}^{N_Z} X_i^{(1)} \mid N_Z = j\right)P\{N_Z = j\} = E(X^{(1)})e^{-\lambda_{NZ}} \sum_{j=0}^{k-1} \frac{(\lambda_{NZ})^j}{j!}$$

We will use formula 5.24.3 from [1]

$$\sum_{n=0}^{k} \frac{1}{n!} x^n = \frac{1}{k!} e^x \Gamma(k + 1, x), \quad (A1)$$

where $\Gamma(a, z)$ is incomplete gamma function defined as

$$\Gamma(a, z) = \int_z^{\infty} e^{-t} t^{a-1} dt.$$
Then first sum in the expression for expectation of total gain is

\[
\sum_{j=0}^{k-1} E \left( \sum_{i=1}^{N_z} X_i^{(1)} \middle| N_z = j \right) P\{N_z = j\} = E[X^{(1)}] \lambda_{N_z} \frac{\Gamma(k-1, \lambda_{N_z})}{\Gamma(k-1)}.
\]

Let us find the second sum

\[
\sum_{j=k}^{\infty} E \left( \sum_{i=1}^{k-1} X_i^{(1)} \middle| N_z = j \right) P\{N_z = j\} = \sum_{j=k}^{\infty} E \left( (k-1)X^{(1)} \right) P\{N_z = j\} =
\]

\[
= (k - 1)E(X^{(1)}) \sum_{j=k}^{\infty} P\{N_z = j\} = (k - 1)E[X^{(1)}] e^{-\lambda_{N_z}} \sum_{j=k}^{\infty} \frac{(\lambda_{N_z})^j}{j!}.
\]

We can find this sum using formula (A1)

\[
e^{-\lambda_{N_z}} \sum_{j=k}^{\infty} \frac{(\lambda_{N_z})^j}{j!} = e^{-\lambda_{N_z}} \left[ \sum_{j=0}^{\infty} \frac{(\lambda_{N_z})^j}{j!} - \sum_{j=0}^{k-1} \frac{(\lambda_{N_z})^j}{j!} \right] =
\]

\[
= e^{-\lambda_{N_z}} \left[ e^{\lambda_{N_z}} - \frac{1}{(k - 1)!} e^{\lambda_{N_z}} \Gamma(k, \lambda_{N_z}) \right] = 1 - \frac{\Gamma(k, \lambda_{N_z})}{\Gamma(k)}.
\]

So the second sum is equal to

\[
\sum_{j=k}^{\infty} E \left( \sum_{i=1}^{k-1} X_i^{(1)} \middle| N_z = j \right) P\{N_z = j\} = (k - 1)E(X^{(1)}) \left( 1 - \frac{\Gamma(k, \lambda_{N_z})}{\Gamma(k)} \right).
\]

The last sum in our formula can be represented as

\[
\sum_{j=k}^{\infty} E \left( \sum_{i=1}^{N_z} X_i^{(2)} \middle| N_z = j \right) P\{N_z = j\} = \sum_{j=k}^{\infty} E \left( \sum_{i=1}^{j} X_i^{(2)} \right) P\{N_z = j\} =
\]
Hence the expectation of total gain is equal to

\[ E(Z) = E\left( \sum_{i=1}^{N_Z} X_i \right) = \]

\[ = E(X^{(1)}) \cdot \lambda_{N_Z} \cdot \left( 1 - \frac{\Gamma(k - 1, \lambda_{N_Z})}{\Gamma(k - 1)} \right) + (k - 1) \cdot \left( 1 - \frac{\Gamma(k, \lambda_{N_Z})}{\Gamma(k)} \right) \]

\[ + E(X^{(2)}) \cdot \lambda_{N_Z} \cdot \left( 1 - \frac{\Gamma(k - 1, \lambda_{N_Z})}{\Gamma(k - 1)} \right) + (1 - k) \cdot \left( 1 - \frac{\Gamma(k, \lambda_{N_Z})}{\Gamma(k)} \right) \]

Hence the expectation of total gain is equal to

\[ E(Z) = E\left( \sum_{i=1}^{N_Z} X_i \right) = \]

\[ = E(X^{(1)}) \cdot \lambda_{N_Z} \cdot \left( 1 - \frac{\Gamma(k - 1, \lambda_{N_Z})}{\Gamma(k - 1)} \right) + (k - 1) \cdot \left( 1 - \frac{\Gamma(k, \lambda_{N_Z})}{\Gamma(k)} \right) \]

\[ + E(X^{(2)}) \cdot \lambda_{N_Z} \cdot \left( 1 - \frac{\Gamma(k - 1, \lambda_{N_Z})}{\Gamma(k - 1)} \right) + (1 - k) \cdot \left( 1 - \frac{\Gamma(k, \lambda_{N_Z})}{\Gamma(k)} \right) \]

Proof of Theorem 4.

\( S_i \) – an experience function on step \( i \) – is a random variable of number consequently correctly recognized \( Q \)-events,

\[ S_0 = 0, \]

\[ S_i = \begin{cases} S_{i-1} + 1, & \text{if } a \text{ occurred,} \\ 0, & \text{if } \bar{a} \text{ occurred (} a \text{ did not occur),} \end{cases} \]

The experience function on step \( i \) can take values from 0 to \( i \) with some probabilities.
We have to know the probabilities $P[S_i < k]$ and $P[S_i \geq k]$, because now
the random variable of single winning $X_i$ takes values $-d, -b, a, c$ with
probabilities

$$\Pr\{X_i = x\} = \begin{cases} p_R q_2, & \text{if } x = -d, \\ p_Q[q_1 \Pr\{S_i < k\} + q_i^* \Pr\{S_i \geq k\}], & \text{if } x = -b, \\ p_Q[p_1 \Pr\{S_i < k\} + p_i^* \Pr\{S_i \geq k\}], & \text{if } x = a, \\ p_R p_2, & \text{if } x = c. \end{cases}$$

Because the probability $P\{S_i < k\}$ is equal to

$$P\{S_i < k\} = 1 - P\{S_i \geq k\} =$$

$$= P\{S_i = 0 \text{ or } S_i = 1 \text{ or } S_i = 2 \text{ or } \ldots \text{ or } S_i = k - 1\} =$$

$$= P\{S_i = 0\} + P\{S_i = 1\} + P\{S_i = 2\} + \ldots + P\{S_i = k - 1\},$$

we will describe all terms.

$P\{S_i = 0\}$ is a probability of experience function to get value 0 on step $i$,
i.e. the device incorrectly recognized coming event $Q$ or it was event $R$. Hence,

$$P\{S_i = 0\} = P\{X_i = -b \text{ or } X_i = c \text{ or } X_i = -d\} =$$

$$= p_Q[q_1 P\{S_{i-1} < k\} + q_i^* P\{S_{i-1} \geq k\}] + p_R p_2 + p_R q_2 =$$

$$= p_Q[q_1 P\{S_{i-1} < k\} + q_i^* P\{S_{i-1} \geq k\}] + p_R.$$

$P\{S_i = 1\}$ is a probability of experience function to get value 1 on step $i$,
i.e. the device correctly recognized coming event $Q$ and mistakes on step $i - 1$. It
means that $S_{i-1} = 0$ and then happens $X_i = a$ (it can be with probability $p_Q p_1$)

$$P\{S_i = 1\} = [p_Q[q_1 P\{S_{i-2} < k\} + q_i^* P\{S_{i-2} \geq k\}] + p_R](p_Q p_1).$$

In the same way we can find another probabilities

$$P\{S_i = 0\}, P\{S_i = 1\}, \ldots, P\{S_i = k - 1\}.$$

For example
\[ P\{S_i = k - 1\} = \]

\[ = [p_Q \left(q_1 P\{S_{i-1-(k-1)} < k\} + q_1^* P\{S_{i-k} \geq k\}\right] + p_R(p_Q p_1)^{k-1}. \]

Now we can evaluate a sum \( P\{S_i < k\} = P\{S_i = 0\} + P\{S_i = 1\} + \]

\[ + P\{S_i = 2\} + \cdots + P\{S_i = k - 1\} \]

taking \( P\{S_j < k\} = 1 - P\{S_j \geq k\}: \)

\[ P\{S_i < k\} = P\{S_i = 0\} + P\{S_i = 1\} + \cdots + P\{S_i = k - 1\} = \]

\[ = p_Q \sum_{j=0}^{k-1} [q_1 P\{S_{i-1-j} < k\} + q_1^* (1 - P\{S_{i-1-j} < k\})](p_Q p_1)^j + \]

\[ + p_R \sum_{j=0}^{k-1} (p_Q p_1)^j = \]

\[ = p_Q(q_1 - q_1^*) \sum_{j=0}^{k-1} P\{S_{i-1-j} < k\}(p_Q p_1)^j + (p_Q q_1^* + p_R) \sum_{j=0}^{k-1} (p_Q p_1)^j. \]

The last sum is geometric progression and the answer is

\[ P\{S_i < k\} = \]

\[ = p_Q(q_1 - q_1^*) \sum_{j=0}^{k-1} P\{S_{i-1-j} < k\}(p_Q p_1)^j + (p_Q q_1^* + p_R) \frac{(p_Q p_1)^{k-1}}{p_Q p_1 - 1}. \]

Using the known dependences between the probabilities, we can express

\[ p_Q p_1 - 1 = p_Q p_1 + p_Q q_1 - p_Q q_1 - 1 = p_Q - p_Q q_1 - 1 = -p_R - p_Q q_1. \]

Finally the probability \( P\{S_i < k\} \) is

\[ P\{S_i < k\} = \]

\[ = p_Q(q_1 - q_1^*) \sum_{j=0}^{k-1} P\{S_{i-1-j} < k\}(p_Q p_1)^j + \frac{p_Q q_1^* + p_R}{p_Q q_1 + p_R} \left(1 - (p_Q p_1)^k\right). (A2) \]
Proof of Theorem 4.

Let us denote \( P\{S_i < k\} = z_i, p_0(q_1 - q_1^*) = \alpha, 0 < \alpha < 1, p_0p_1 = \beta, 0 \leq \beta < 1, \frac{p_0q_1 + p_R}{p_0q_1 + p_R} \left(1 - (p_0p_1)^k\right) = \gamma, 0 < \gamma < 1, \) and \( \gamma = \frac{(1-\beta^k)(1-\alpha-\beta)}{(1-\beta)} \), then the equation (A2) is written as

\[
z_i = \alpha \cdot \sum_{s=i-k}^{i-1} \beta^{i-1-s} \cdot z_s + \gamma
\]

or

\[
z_i = \alpha \cdot (z_{i-1} + \beta z_{i-2} + \beta^2 z_{i-2} + \cdots + \beta^{k-1} z_{i-k}) + \gamma.
\]

As all \( z_i = P\{S_i < k\} \) are probabilities and \( \forall i \ 0 \leq z_i \leq 1 \), the sequence \( \{z_i\}_{i=1}^{\infty} \) is bounded from both sides.

This sequence is nonincreasing, we will prove it by induction. The first \( k \) terms are equal to 1: \( z_0 = z_1 = z_2 = \cdots = z_{k-1} = 1 \), because \( S_0 = 0 \) and if \( i < k \) then \( z_i = P\{S_i < k\} = 1 \).

First of all we prove \( z_k \leq z_{k-1} \) to have \( k \) inequalities for the induction

\[
z_k = \alpha(1 + \beta + \beta^2 + \cdots + \beta^{k-1}) + \gamma = \alpha \frac{(1-\beta^k)}{(1-\beta)} + \frac{(1-\beta^k)(1-\alpha-\beta)}{(1-\beta)} = 1 - \beta^k \leq 1 = z_{k-1}.
\]

Let inequality \( z_{i-1} \leq z_{i-2} \) holds for all \( i - 1 \) first members of the sequence. We will take \( z_i \) and prove that \( z_i \leq z_{i-1} \). The difference between two terms of the sequence is

\[
z_i = \alpha \cdot (z_{i-1} + \beta z_{i-2} + \beta^2 z_{i-2} + \cdots + \beta^{k-1} z_{i-k}) + \gamma
\]

\[
z_{i-1} = \alpha \cdot (z_{i-2} + \beta z_{i-3} + \beta^2 z_{i-4} + \cdots + \beta^{k-1} z_{i-k-1}) + \gamma
\]

\[
z_i - z_{i-1} = \alpha \cdot (z_{i-1} + (\beta - 1)z_{i-2} + \beta(\beta - 1)z_{i-2} + \cdots + \beta^{k-2}(\beta - 1)z_{i-k} - \beta^{k-1}z_{i-k-1}) \leq \]

31
\[
\leq \alpha \cdot (z_{i-2} + (\beta - 1)z_{i-2} + \beta(\beta - 1)z_{i-2} + \cdots + \beta^{k-2}(\beta - 1)z_{i-k} \\
- \beta^{k-1}z_{i-k-1}) = \\
= \alpha \cdot (\beta z_{i-2} + \beta(\beta - 1)z_{i-2} + \cdots + \beta^{k-2}(\beta - 1)z_{i-k} - \beta^{k-1}z_{i-k-1}) = \\
= \alpha \cdot (\beta^2 z_{i-2} + \beta^2(\beta - 1)z_{i-3} + \cdots + \beta^{k-2}(\beta - 1)z_{i-k} - \beta^{k-1}z_{i-k-1}) \leq \cdots \leq \\
\leq \alpha(\beta^{k-1}z_{i-k} - \beta^{k-1}z_{i-k-1}) = \alpha \beta^{k-1}(z_{i-k} - z_{i-k-1}) \leq 0,
\]

because of \( 0 < \alpha < 1, 0 \leq \beta < 1, \) and \( z_{i-k} \leq z_{i-k-1}. \)

So \( z_i \leq z_{i-1} \) and the sequence \( \{z_i\}_{i=1}^\infty \) is monotonic (nonincreasing). If nonincreasing sequence is bounded from below, then this sequence converges [4], hence, the sequence \( \{z_i\}_{i=1}^\infty \) has a limit \( \lim_{i \to \infty} z_i = z. \)

Let \( i \to \infty, \) then

\[
z = \lim_{i \to \infty} z_i = \lim_{i \to \infty} \left[ \alpha \sum_{s=i-k}^{i-1} \beta^{i-1-s}z_s + \gamma \right] = \alpha \sum_{s=i-k}^{i-1} \beta^{i-1-s}z_s + \gamma = z\alpha \frac{1 - \beta^k}{1 - \beta} + \gamma
\]

i.e.

\[
z = \frac{(1 - \beta)\gamma}{1 - \beta - \alpha(1 - \beta^k)}.
\]

Rewriting it in the initial notations, we obtain

\[
\lim_{i \to \infty} P\{S_i < k\} = \frac{(p_0 q_i^* + p_R)(1 - (p_0 p_1)^k)}{(p_0 q_1 + p_R) - p_Q(q_1 - q_1^*)(1 - (p_0 p_1)^k)}.
\]

\[\blacksquare\]
Appendix 2

1. Evaluations with the closing prices

In the Section 4 we used the average of the opening and closing prices as a daily price (Table 5), but we can do the same analysis using only closing prices (Table 6).

Table 4. The model parameters evaluated by the average price, %

<table>
<thead>
<tr>
<th>Threshold</th>
<th>( \lambda )</th>
<th>( \mu )</th>
<th>( a )</th>
<th>( -b )</th>
<th>( c )</th>
<th>( -d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>6%</td>
<td>246</td>
<td>4,61</td>
<td>-0,64</td>
<td>2,81</td>
<td>-2,93</td>
<td></td>
</tr>
<tr>
<td>8%</td>
<td>248</td>
<td>2,63</td>
<td>-0,65</td>
<td>4,00</td>
<td>-2,80</td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>249</td>
<td>1,63</td>
<td>-0,66</td>
<td>3,74</td>
<td>-3,05</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. The model parameters evaluated by the closing price, %

<table>
<thead>
<tr>
<th>Threshold</th>
<th>( \lambda )</th>
<th>( \mu )</th>
<th>( a )</th>
<th>( -b )</th>
<th>( c )</th>
<th>( -d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>6%</td>
<td>246</td>
<td>0,85</td>
<td>-0,90</td>
<td>3,92</td>
<td>-3,65</td>
<td></td>
</tr>
<tr>
<td>8%</td>
<td>248</td>
<td>0,88</td>
<td>-0,92</td>
<td>5,59</td>
<td>-3,90</td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>249</td>
<td>0,89</td>
<td>-0,93</td>
<td>5,47</td>
<td>-3,87</td>
<td></td>
</tr>
</tbody>
</table>

2. Evaluations with the return of the index

Of course, volatility is not the only indicator of market behavior and a signal of crisis. The question of the calculation period naturally rises in estimation of the volatility. If we shorten the period, errors in the evaluation of volatility will grow because it will get a higher variability. If we make the period longer, it will reduce errors but will delay the information observed about the market. Perhaps it is also the case in our calculations.
We carry out the same calculations as in Section 4 for estimation of the model parameters, using the index returns instead of the index volatility

\[ r_i = \frac{S_i - S_{i-1}}{S_{i-1}}, i = 2, \ldots, 2665. \]

The fall of the index immediately reflected in its return (Fig. 12), or more precisely on its amplitude. The estimates of all parameters are given in Table 6.

![Index S&P 500 and its return](image)

**Fig. 12. Index S&P 500 and its return**

**Table 6. Estimates of parameters**

<table>
<thead>
<tr>
<th>The threshold value</th>
<th>( \lambda )</th>
<th>( \mu )</th>
<th>( a )</th>
<th>( -b )</th>
<th>( c )</th>
<th>( -d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3%</td>
<td>246</td>
<td>4</td>
<td>7</td>
<td>-7</td>
<td>40</td>
<td>-39</td>
</tr>
<tr>
<td>4%</td>
<td>248</td>
<td>2</td>
<td>7</td>
<td>-7</td>
<td>45</td>
<td>-50</td>
</tr>
<tr>
<td>5%</td>
<td>249</td>
<td>1</td>
<td>7</td>
<td>-8</td>
<td>-47</td>
<td>-51</td>
</tr>
</tbody>
</table>

So, the return can also serve as a ‘yardstick’ of the index state and give even more adequate assessment for our model.
In addition, the use of average ratings for such a long period of time inevitably oversimplifies the calculations. It is interesting to look at the distribution of crisis events by years (Table 7). The lack of numbers means the absence of volatility ‘jumps’ more than 6% in a specified period, i.e. ‘quiet life’ and the absence of shocks in the market.

*Table 7. Estimates of model parameters with 6% threshold by years*

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>8</td>
<td>11</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>b</td>
<td>-9</td>
<td>-11</td>
<td>-9</td>
<td>-8</td>
<td>-5</td>
<td>-4</td>
<td>-4</td>
<td>-4</td>
<td>-7</td>
<td>-12</td>
<td>-9</td>
</tr>
<tr>
<td>c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td>14</td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-30</td>
<td>-15</td>
</tr>
</tbody>
</table>

3. Defining the threshold

We estimated $a$ and $b$ at points corresponding to the event $Q$ using the threshold value: if the index goes up at this moment (the volatility is positive and less than the threshold), it means that the event $a$ was realized, and if it goes down – then $–b$ is observed. The same approach was used for $R$-events, i.e. if the index has increased, then the change of the index is $c$; if the index has fallen, the change is $–d$. But there can be questions about the threshold – why do we use the 6% volatility as the reason to define $Q$ and $R$ events?

You can see the variational series and the histogram built on volatility data.
Fig. 13. Variational series of the volatility data

Fig. 14. Histogram of the volatility series

Values from the tail of variational series and from histogram correspond to the crises days and the greater part of values corresponds to the ‘regular’ days. 6% threshold was chosen to distinguish them. You can use another threshold to identify various crises – the global crisis has greater volatility than the local crisis or the crisis in some sector.

Also we estimated all parameters for the threshold volatility from 3% to 10% and you can see on Fig. 15 the estimates of $a$ and $-b$ change a little, but $c$ and $-d$ grow when the threshold increases. This happens because the threshold influence to the magnitude of the crises, also the low threshold value will cause overestimation of the number of crises days and $c, d$ become smaller. We choose
as appropriate the threshold of 6% to ‘catch the big crises’ and not to affect the smaller (Fig. 5).

Fig. 15. Estimates of the parameters by the volatility. The threshold is on the axis OX

The same picture can be drawn for the return method with the average (Fig. 16) and closing prices (Fig. 17).

Fig. 16. Estimates of the parameters by the return of the average index prices. The threshold is on the axis OX
4. Top 10 companies analysis

We qualify $Q$- and $R$-events using the S&P 500 data, as it was done in Section 4, and then evaluate single winnings $a, b, c, d$ for all 10 firms: if its stock goes up/down when the volatility of S&P 500 is less than 6% than it is a realization of $a/b$; and if its stock goes up/down when the volatility of S&P 500 is greater than 6% than it is a realization of $c/d$.

The estimates you can see on the Table 3. The first row shows the parameters of S&P 500 and filled rows show what companies have stocks more volatile than the stock index S&P 500 (if all parameters in absolute value are greater than the same for S&P 500).

Table 8. Estimates with the average prices

<table>
<thead>
<tr>
<th>Company</th>
<th>threshold 6%, lambda=246, mu=4</th>
<th>threshold 8%, lambda=248, mu=2</th>
<th>threshold 10%, lambda=249, mu=1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a$ $-b$ $c$ $-d$</td>
<td>$a$ $-b$ $c$ $-d$</td>
<td>$a$ $-b$ $c$ $-d$</td>
</tr>
<tr>
<td>S&amp;P500</td>
<td>0.6 -0.6 2.8 -2.9</td>
<td>0.6 -0.7 4.0 -2.8</td>
<td>0.6 -0.7 3.7 -3.1</td>
</tr>
<tr>
<td>Exxon</td>
<td>0.8 -0.9 3.1 -3.8</td>
<td>0.9 -0.9 4.1 -5.5</td>
<td>0.9 -0.9 4.7 -5.7</td>
</tr>
<tr>
<td>Microsoft</td>
<td>1.1 -1.1 3.0 -3.3</td>
<td>1.2 -1.1 3.5 -3.1</td>
<td>1.2 -1.1 4.0 -3.7</td>
</tr>
</tbody>
</table>
Table 9. Estimates with the closing prices

<table>
<thead>
<tr>
<th>Company</th>
<th>threshold 6%, lambda=246, mu=4</th>
<th>threshold 8%, lambda=248, mu=2</th>
<th>threshold 10%, lambda=249, mu=1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a$</td>
<td>$-b$</td>
<td>$c$</td>
</tr>
<tr>
<td>S&amp;P500</td>
<td>0.9</td>
<td>-0.9</td>
<td>3.9</td>
</tr>
<tr>
<td>Exxon</td>
<td>1.2</td>
<td>-1.2</td>
<td>5.6</td>
</tr>
<tr>
<td>Microsoft</td>
<td>1.5</td>
<td>-1.4</td>
<td>5.7</td>
</tr>
<tr>
<td>General Electric</td>
<td>1.4</td>
<td>-1.4</td>
<td>6.2</td>
</tr>
<tr>
<td>Morgan Chase</td>
<td>1.9</td>
<td>-1.7</td>
<td>8.9</td>
</tr>
<tr>
<td>Proctor&amp;Gamble</td>
<td>1.0</td>
<td>-1.0</td>
<td>3.2</td>
</tr>
<tr>
<td>Johnson&amp;Johnson</td>
<td>0.9</td>
<td>-0.9</td>
<td>3.1</td>
</tr>
<tr>
<td>Apple</td>
<td>2.3</td>
<td>-2.2</td>
<td>4.9</td>
</tr>
<tr>
<td>AT&amp;T</td>
<td>1.4</td>
<td>-1.3</td>
<td>4.2</td>
</tr>
<tr>
<td>IBM</td>
<td>1.3</td>
<td>-1.2</td>
<td>3.1</td>
</tr>
<tr>
<td>Bank of America</td>
<td>1.8</td>
<td>-1.7</td>
<td>8.1</td>
</tr>
</tbody>
</table>
References