Review of Business and Economics Studies

REVIEW OF BUSINESS AND ECONOMICS STUDIES (ROBES) is the quarterly peer-reviewed scholarly journal published by the Financial University under the Government of Russian Federation, Moscow. Journal's mission is to provide scientific perspective on wide range of topical economic and business subjects.

CONTACT INFORMATION
Financial University
Leningradsky prospekt, 53, office 5.6
123995 Moscow
Russian Federation
Telephone: +7 (499) 943-98-02
Website: www.robes.fa.ru

AUTHOR INQUIRIES
Inquiries relating to the submission of articles can be sent by electronic mail to robes@fa.ru.

ISSN 2308-944X

Вестник исследований бизнеса и экономики

Редакция научных журналов Финансового университета
123995, Москва, ГСП-5, Ленинградский пр-т, 53, комн. 5.6
Тел. 8 (499) 943-98-02.
Интернет: www.robes.fa.ru.

Журнал “Review of Business and Economics Studies” («Вестник исследований бизнеса и экономики») зарегистрирован в Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций 15 сентября 2016 г. Свидетельство о регистрации ПИ № ФС77-67072.

Подписано в печать: **14.03.2018.**
Формат **60 × 84 1/8.**
Заказ № 238 от 14.03.2018.
Отпечатано в Отделе полиграфии Финуниверситета (Ленинградский проспект, д. 49).
16+

Review of Business and Economics Studies

Volume 6, Number 1, 2018

CONTENTS

Optimization of the Shape of the Pareto Set in the Problems of Multi-criterial Programming

Bula Katendi Axel, E.A. Umnov, A.E. Umnov **5**

Does Enterprise Value Really Depend on WACC and Free Cash Flow? The Evidence of Irrationality from the Oil and Gas Sector

Pavel E. Zhukov **17**

Weather Derivatives in Russia: Farmers' Insurance against Temperature Fluctuations

Eric Carkin, Stanislav Chekirov, Anastasia Echimova, Caroline Johnston, Congshan Li, Vladislav Secrieru, Alyona Strelnikova, Marshall Trier, Vladislav Trubnikov **29**

Does Social Inequality Stimulate the Economic Growth? (On the examples of the chosen developing countries)

Alina Pukhaeva, Elena Miroshina (Silantieva) **43**

Country Risk in International Investment Its' Structure and Methods of Estimation

Alexey Ivkin **56**

Business Valuation of Nike

Elena Miroshina (Silantieva), Egor Romanov **78**

"Uniform Subsidy" and New Trends in Financing of Agricultural Insurance in Russian Federation

L.Yu. Piterskaya, N.A. Tlisheva, A.V. Piterskaya **83**

Вестник исследований бизнеса и экономики
№ 1, 2018

СОДЕРЖАНИЕ

Оптимизация формы множества Парето в задачах многокритериального программирования
Була Аксел Катенди, Е.А. Умнов, А.Е. Умнов 5

Разве стоимость компании действительно зависит от средневзвешенной стоимости капитала и свободного денежного потока? Свидетельства иррациональности в нефтегазовом секторе
Павел Е. Жуков 17

Погодные деривативы в России: страхование фермеров от колебаний температуры
Eric Carkin, Станислав Чекиров, Анастасия Екимова, Caroline Johnston, Congshan Li, Vladislav Secrieru, Алена Стрельникова, Marshall Trier, Владислав Трубников 29

Стимулирует ли социальное неравенство экономический рост? (на примерах выбранных развивающихся стран)
Алина Пухаева, Елена Мирошина (Силантьева) 43

Страновой риск в международных инвестициях: структура и методы расчета
Алексей Ивкин 56

Оценка бизнеса корпорации Nike
Елена Мирошина (Силантьева), Егор Романов 78

«Единая субсидия» и новые веяния в субсидировании агрострахования в Российской Федерации
Людмила Питерская, Нафсэт Тлишева, Анастасия Питерская 83

Review of Business and Economics Studies
2018, Vol. 6, No. 1, 5–16
DOI: 10.26794/2308-944X-2018-6-1-5-16

Optimization of the Shape of the Pareto Set in the Problems of Multi-criterial Programming

Bula Katendi Axel[1], E.A. Umnov[2], A.E. Umnov[2]
Moscow Institute of Physics and Technology (State University)
[1]axelbula@gmail.com, [2]mail@umnov.ru

Abstract
In this paper, a scheme for using the method of smooth penalty functions for the dependence of solutions of multi-criterial optimization problems on parameters is being considered. In particular, algorithms based on the method of smooth penalty functions are given to solve problems of optimization by the parameters of the level of consistency of the objective functions and to find the corresponding shape of the Pareto's set.
Keywords: multi-criterial parametrical programming tasks; set of Pareto; method of smooth penalty functions; optimization problem in terms of parameters.
JEL classification: C61, C65

In mathematical modeling, it is often necessary to formalize preferences for states of the modeled object that generates several independent target functions. According to the historically established tradition, in this case, it is customary to talk about of multi-criterial optimization problems.

A finite-dimensional multi-criterial model is a mathematical model with N objective functions:

$$f_k(x,u) \to \max_x \quad k = [1, N], \tag{1}$$

subject to maximization possessing at interior points of the set of elements $x \in E^n$, and satisfying the following conditions:

$$y_i(x,u) \le 0 \quad i = [1, m] \tag{2}$$

where $u \in \Theta \subseteq E^r$ — vector of parameters of the model. It is assumed that the functions $f_k(x,u)$ and $y_i(x,u)$ are sufficiently smooth, i.e. they have continuous derivatives of a desired order in all their arguments.

The incorrectness in the general case of such a statement is obvious, since the element x that is extremal for one of objective functions, in general, is not such for others.

However, useful information can be obtained by successively solving the following problems with a criterion for finding an extremum on the set (2) of each of the functions (1) separately for $k = [1, N]$:

$$f_k(x,u) \to \max_x$$

$$\text{subject to } y_i(x,u) \le 0 \quad i = [1, m] \tag{3}$$

The objective function $f_k(x,u)$ is called improvable in the feasible point x_0 (i.e., satisfying the condition (3)) if there is another feasible point x_1, for which $f_k(x_1,u) > f_k(x_0,u)$.

It is clear that the solution of problem (3) for any $k = [1, N]$ is un-improvable, or "non-ideal" at the point of view of the other objective function $f_k(x,u)$.

The concept of improving the multi-criterial objective function allows the feasible points to be divided into two subsets: for the first, all feasible points improve all objective functions and for the second, there are points for which the improvement of one function causes the deterioration of at least one other function.

The second subset is called a Pareto-type set or, simply, a Pareto set.

A general universal approach to the solution of multi-criterial optimization problems has not been proposed yet, but numerous approaches have been developed (see Fiacco & McCormick, 1968; Lotov & Pospelov, 2008), which limit the number of solutions.

For example, in the practical use of multi-criterial mathematical models, the set of independent objective functions is often replaced by a single one, thus passing to the standard problem of mathematical programming, allowing finding consistent or compromising solutions on the Pareto set in a certain sense.

Statement of the problem

In this article, the problem of finding an element on the set (2) that minimizes the gap between the objective functions will be considered as a compromise. In other words, this is a mathematical programming problem of the following form:

$$\begin{aligned} &\text{minimize } \rho\text{, subject to } \rho \geqslant 0, \\ &y_i(x,u) \le 0 \qquad i = [1,m], \\ &f_k(x,u) \ge F_k^*(u) - \rho \qquad k = [1,N] \end{aligned} \tag{4}$$

whose solution will be denoted by $\rho^{**}(u)$ and $x^{**}(u)$. Here $F_k^*(u) = f_k\left(x_k^*(u),u\right)$ and ρ is called "mismatch value".

The problem (4) is naturally called a two-level parametric problem since in its formulation it contains solutions $F_k^*(u)$ $k = [1,N]$ of problems with single criterion (3), which we call first level problems. In this case, both in the problems of the first and the second level, it is assumed that the vector of the parameters $u \in \Theta$ is fixed.

It is clear that the extreme value of the mismatch between the criteria in the general case is determined by the properties of the Pareto set and depends on the parameters vector u. Therefore, it is natural to indicate the third level optimization problem for the models (1)–(2) as follows:

$$\text{optimize the expression } \rho^{**}(u) \text{ subject to } u \in \Theta \tag{5}$$

which solution will be the vector of parameters $u^{***} \in \Theta$ and the number $\rho^{***} = \rho^{**}\left(u^{***}\right)$. In the present paper, possible solutions to problem (5) will be considered.

Solution method

Let us consider the problem of finding in the parameter space a standard method (for example, gradient) of finding the extremum of the mismatch value of the objective functions of the multi-criterial model (3)–(4)–(5).

The specificity of this problem is based on the fact that the formulation of the problem (5) (the upper level or third level) includes the dependence $\rho^{**}(u)$, the solution of the problem (4) (the second level) which in turn, depends on $F_k^*(u) = f_k\left(x_k^*(u),u\right)$ $\forall k = [1,N]$ — the solutions of the problem (3) (the lower level or first level).

The functions $\rho^{**}(u)$ and $F_k^*(u)$ $\forall k = [1,N]$ in the general case (even for smooth functions $f_k(x,u)$ and $y_i(x,u)$) may be not differentiable, that why the use of any numerical method based on Taylor approximations is not possible.

It is proposed the use of the method of smooth penalty function to overcome this difficulty (see Umnov, 1975) and obtain a sufficiently smooth approximation dependences of $\rho^{**}(u)$ and $F_k^*(u)$ $\forall k=[1,N]$.

It's assumed that the penalty function $P(\tau,s)$, which penalizes the restriction $s\le 0$, satisfies the following conditions:

1 $\forall\tau>0$ and $\forall s$, the function $P(\tau,s)$ has continuous derivatives with respect to all its arguments up to the second order

2 $\forall\tau>0$ and $\forall s$,

$$\frac{\partial P}{\partial s}>0 \;\; ; \;\; \frac{\partial^2 P}{\partial s^2}>0 \tag{6}$$

3 $P(\tau,s)>0\,\forall s$ and $\forall\tau>0$, and,

$$\lim_{\tau\to+0} P(\tau,s)=\begin{cases}+\infty, & s>0\\ 0, & s<0\end{cases} \tag{7}$$

When solving the third-level problem by an iterative method, for each step of the method, it is necessary preliminary to solve the problems of the second and first levels for a fixed vector of parameters u. Let us first consider a possible scheme for solving first-level problems. In fact, we will use an auxiliary function for the one-criterion problems (3), as follows:

$$A_k(\tau,x,u)=f_k(x,u)-\sum_{i=1}^{m}P(\tau,y_i(x,u)) \qquad \forall k\in[1,N] \tag{8}$$

while a sufficiently smooth penalty function $P(\tau,s)$ satisfies conditions (6) and (7).

As shown in Zhadan (2014), instead of the smooth approximations $x_k^*(u)$ solutions of each task of problem (3), we can take $\overline{x}_k(u)$ stationary points of the auxiliary function (8), defined like:

$$\frac{\partial A_k}{\partial x_j}=0 \qquad \forall j\in[1,n] \tag{9}$$

or

$$\frac{\partial f_k}{\partial x_j}-\sum_{i=1}^{m}\frac{\partial P}{\partial y_i}\frac{\partial y_i}{\partial x_j}=0 \qquad \forall j\in[1,n]$$

Since the condition of the second-level problem (4) includes the dependencies $F_k^*(u)=f_k(x_k^*(u),u)$ $\forall k=[1,N]$ which are not differentiable functions for all their arguments, then for these dependencies it is also necessary to choose a smoothed approximation.

As an approximation, the auxiliary function calculated at a stationary point $\overline{F}_k(u)=A_k(\tau,\overline{x}_k(u),u)$ can be used, because (due to the properties of the penalty function method) its value for small positive τ is close to the optimal value of the objective function of the k-th problem (3).

Standard optimization methods used for lower-level tasks, based on the use of continuous gradients or other differential characteristics, suggest that in addition to the solving system (9), these characteristics themselves can be found.

Let us demonstrate this using the example of calculating the derivatives of the function $\overline{F}_k(u)$ with respect to the components of the vector u of parameters.

As $\overline{F}_k(u)=A_k(\tau,\overline{x}_k(u),u)$, then according to the rule for differentiating a composite functionof several variables, we have:

$$\frac{\partial \overline{F}_k}{\partial u_p} = \frac{\partial A_k}{\partial u_p} + \sum_{j=1}^{n} \frac{\partial A_k}{\partial x_j}\frac{\partial x_j}{\partial u_p} \qquad \forall p \in [1,r]$$

Using (9), we have:

$$\frac{\partial \overline{F}_k}{\partial u_p} = \frac{\partial A_k}{\partial u_p}\left(\tau, \overline{x}_k(u), u\right) \qquad \forall p \in [1,r] \tag{10}$$

Note that the last simplification would be impossible if for $F_k^*(u)$ a more natural approximation $f_k\left(x_k^*(u),u\right)$ is used instead of the smoothing approximation $A_k\left(\tau,\overline{x}_k(u),u\right)$.

Let us now look into the solution to the second-level problem. To make application of the penalty function method more convenient, the problem (4) is expressed as:
maximize $-\rho$, subject to $-\rho \geqslant 0$,

$$y_i(x,u) \le 0 \qquad i = [1,m], \tag{11}$$

$$Y_k(\rho,x,u) \le 0 \qquad k = [1,N]$$

where $Y_k(\rho,x,u) = F_k^*(u) - \rho - f_k(x,u)$

The solution to this problem will be denoted by $\rho^{**}(u)$ and $x^{**}(u)$.

Let us define the auxiliary function for the problem (10) as follows:

$$E(\tau,\rho,x,u) = -\rho - P(\tau,-\rho) - \sum_{k=1}^{N} P\left(\tau, Y_k(\rho,x,u)\right) - \sum_{i=1}^{m} P\left(\tau, y_i(x,u)\right) \tag{12}$$

replacing previously in $Y_k(\rho,x,u)$ the dependency $F_k^*(u)$ by its smoothed approximation $\overline{F}_k(u)$.

For the set of variables $\{\rho, x_1, x_2, \ldots, x_n\}$, the conditions for the stationarity of the auxiliary function (12) will be:

$$\begin{cases} \dfrac{\partial E}{\partial \rho} = -1 + \dfrac{\partial P}{\partial \rho} + \displaystyle\sum_{k=1}^{N} \dfrac{\partial P}{\partial Y_k} = 0 \\ \dfrac{\partial E}{\partial x_j} = \displaystyle\sum_{k=1}^{N} \dfrac{\partial P}{\partial Y_k}\dfrac{\partial f_k}{\partial x_j} - \sum_{i=1}^{m} \dfrac{\partial P}{\partial y_i}\dfrac{\partial y_i}{\partial x_j} = 0 \qquad \forall j = [1,N] \end{cases} \tag{13}$$

Let the solutions of system (13) be $\overline{\overline{\rho}}(u)$ and $\overline{\overline{x}}(u)$, then, as a smoothed approximation of the dependency $\rho^{**}(u)$, we can use the function $\overline{\overline{E}}(u) = -E\left(\rho, \overline{\overline{\rho}}(u), \overline{\overline{x}}(u), u\right)$. The derivatives of this function by the components of the vector u ofparameters and the rule for differentiating a composite function of several variables give us:

$$\frac{\partial \overline{\overline{E}}}{\partial u_p} = \frac{\partial E}{\partial u_p} + \sum_{j=1}^{n} \frac{\partial E}{\partial x_j}\frac{\partial x_j}{\partial u_p} + \frac{\partial E}{\partial \rho}\frac{\partial \rho}{\partial u_p} \qquad \forall p \in [1,r].$$

From (13) we know that $\dfrac{\partial E}{\partial \rho} = 0$ and $\dfrac{\partial E}{\partial x_j} = 0 \ \ \forall j = [1,N]$.

Then the last expression can be written simply:

$$\frac{\partial \overline{\overline{E}}}{\partial u_p} = \frac{\partial E}{\partial u_p}\left(\rho, \overline{\overline{\rho}}(u), \overline{\overline{x}}(u), u\right) \quad \forall p \in [1, r] \tag{14}$$

Finally, we obtain formulas for the gradient components of $\overline{\overline{E}}(u)$ in terms of the functions usedin the formulation of the multicriterial model (4)–(5) and the method of smooth penalty functions. From (12) it is obtained:

$$\frac{\partial \overline{\overline{E}}}{\partial u_p} = -\sum_{k=1}^{N} \frac{\partial P}{\partial Y_k} \frac{\partial Y_k}{\partial u_p} - \sum_{k=1}^{N} \frac{\partial P}{\partial y_i} \frac{\partial y_i}{\partial u_p}$$

where $\frac{\partial Y_k}{\partial u_p} = \frac{\partial \overline{F}_k}{\partial u_p} - \frac{\partial f_k}{\partial u_p}$ and the value of $\frac{\partial \overline{F}_k}{\partial u_p}$ can be found from (10).

Formulas (14) allow us to solve the third-level problem by applying any of the first-order methods, for example, conjugate directions. Note that second-order methods should also be considered here. However, this will be done at the end of the article, while now let us illustrate an example.

Proposed method in use

Let us consider multi-criterial mathematical model in which $x = \| x_1 x_2 x_3 \|^T \in E^3$ is a vector of independent variables and $u = \| u_1 u_2 \|^T \in E^2$ is a vector of parameters.

The problem is to maximize for x and $u \in \Theta$ the functions:

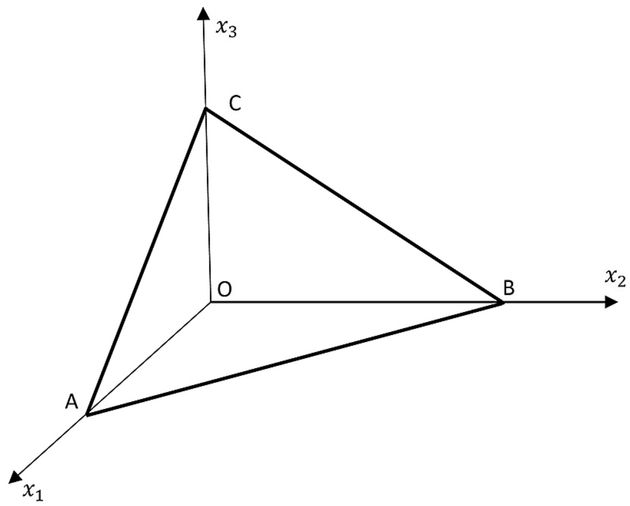


Figure 1. Geometric interpretation of models (6)–(7).

$$f_1(x,u) = x_1, \quad f_2(x,u) = x_2, \quad f_2(x,u) = x_3$$

subject to $x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0,\ a_1(u_1, u_2)x_1 + a_2(u_1, u_2)x_2 + a_3(u_1, u_2)x_3 \leq b(u_1, u_2)$

and where the functions $a_1(u_1, u_2), a_2(u_1, u_2), a_3(u_1, u_2)\, and\, b(u_1, u_2)$ are given by the condition below.

A valid region of the model (with an allowable fixed u) is a rectangular pyramid OABC. The Pareto set coincides with the face of ABC or is a part of it. $A = \| u_1 0 0 \|^T$ and $B = \| 0 u_2 0 \|^T$.

We assume that the set Θ in the parameter space is given by the condition that the sum of the lengths of the segments OA, OB and OC is constant and equal 3.

Applying the standard methods of analytic geometry, we find that for the compatibility of the system of model constraints, the existence of $r \geq 0$ is necessary such that:

$$a_1(u_1, u_2) = u_2 r$$

$$a_2(u_1, u_2) = u_1 r$$

$$a_3(u_1, u_2) = u_1 u_2$$

$$b(u_1, u_2) = u_1 u_2 r$$

By choosing r such that $r = 3 - u_1 - u_2$, we assure that the set Θ will not be empty when $0.1 \leq u_1 \leq 2.5$ and $0.1 \leq u_2 \leq 2.5$.

The minimum value of the discrepancy between the criteria in this example depends on the form of the Pareto set, which is the triangle ABC, or a part of it. A graphic representation of the dependence of the error value of the objective functions on the parameters u_1 and u_2 is shown in pictures 2 and 3. Let us see with more details the properties of this dependence.

Clearly, the solutions of the first-level problems (3) for fixed u_1 and u_2 are:

$$f_1\left(x^*(u)\right) = u_1, \quad f_2\left(x^*(u)\right) = u_2, \quad f_3\left(x^*(u)\right) = 3 - u_1 - u_2.$$

Consequently, the task of the second level (4) — minimizing the discrepancy of the criteria, will have the form:
minimize ρ according to $\{x_1, x_2, x_3, \rho\}$
subject to $\rho \geq 0$,

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0,$$

$$u_2 r x_1 + u_1 r x_2 + u_1 u_2 x_3 \leq u_1 u_2 r \ ,$$

$$x_1 \geq u_1 - \rho,$$

$$x_2 \geq u_2 - \rho,$$

$$x_3 \geq r - \rho \text{'}$$

$$r = 3 - u_1 - u_2$$

His solution will be designed by $\rho^{**}(u_1, u_2)$.

Finally, the task of the third level (5) for our example will be:
minimize $\rho^{**}(u_1, u_2)$ by $\{u_1, u_2\}$ when $0.1 \leq u_1 \leq 2.5$ and $0.1 \leq u_2 \leq 2.5$.

It is known from the theory of mathematical programming that the properties of the dependence $\rho^{**}(u_1, u_2)$ are primarily determined by how the set of constraints of a model of the «inequality» type is divided into active and inactive ones, that is, the first of which are satisfied as equalities, and the second — as strict inequalities.

This separation depends on the values of the parameters of the model and its optimal variant determines the solution of the second-level problem.

First, suppose that the values of the model parameters initiate a conflict of all three criteria simultaneously. In other words, the improvement of the value of any one of the objective functions of the model is possible only if the values of all the others deteriorate.

In this case, the last five constraints of the second-level problem must be active, and we obtain

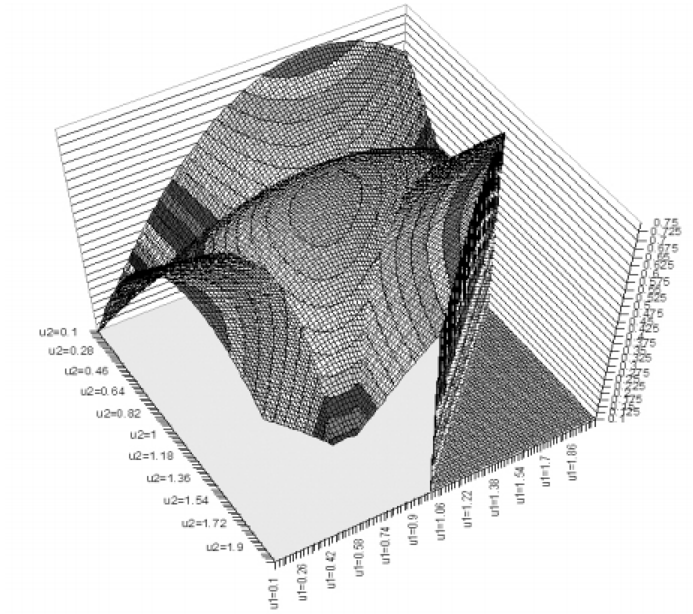


Figure 2. The quantity $|OA|+|OB|+|OC|$ is constant (graphic — 3D).

the following system of equations, which allows us to find the analytical form of the dependence $\rho^{**}(u_1,u_2)$.

$$\begin{cases} u_2rx_1+u_1rx_2+u_1u_2x_3=u_1u_2r \\ x_1=u_1-\rho \\ x_2=u_2-\rho \\ x_3=r-\rho \\ r=3-u_1-u_2 \end{cases}$$

Solving this system above, we have the analytic form of ρ which depends on u_1 and u_2:

$$\rho^{**}(u_1,u_2)=\frac{2}{\dfrac{1}{u_1}+\dfrac{1}{u_2}+\dfrac{1}{3-u_1-u_2}}$$

The stationary points of $\rho^{**}(u_1,u_2)$ are $\|1\ 1\|^T$, $\|-3\ 3\|^T$, $\|3 -3\|^T$ and $\|3\ \ 3\|^T$ can be easily found, and for the first point the function has a local maximum with the value 2/3 according to the Sylvester criterion, not for the others because they don't satisfy the condition of non-negativity of the variables x_1, x_2, x_3 and r.

The formula obtained is valid only in a certain area contained in Θ. An analysis of the isoline system shown in Picture 3, allows selecting five areas with different sets of active restrictions. Light lines determine the boundaries between the areas. The formula obtained above is valid only in area 4. In this area, the Pareto set of the model under consideration consists of the interior points of the triangle ABC.

Outside area 4, the formula for $\rho^{**}(u_1,u_2)$ is different. For the area 1, for example, $\rho^{**}(u_1,u_2)$ is found from the system of equations:

$$\begin{cases} u_2rx_1+u_1rx_2+u_1u_2x_3=u_1u_2r \\ x_1=0 \\ x_2=u_2-\rho \\ x_3=r-\rho \\ r=3-u_1-u_2 \end{cases}$$

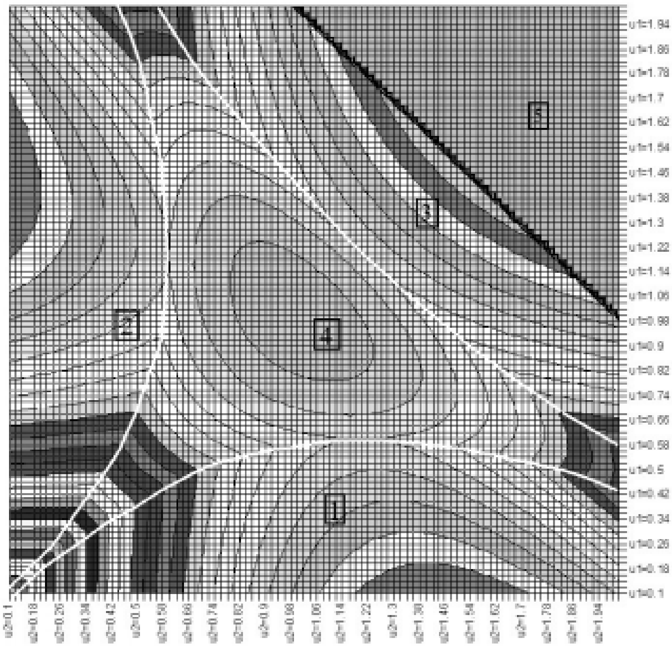


Figure 3. The quantity $|OA|+|OB|+|OC|$ is constant (isoline view).

since the set of active restrictions in it is different: it contains the condition $x_1=0$, instead of $x_1=u_1-\rho$. We can easily prove that in the area 1

$$\rho^{**}(u_1,u_2)=\frac{2}{\dfrac{1}{u_2}+\dfrac{1}{3-u_1-u_2}}$$

There are not stationary points for this dependency.

For areas 2 and 3, the arguments and results are similar. The Pareto sets in areas 1, 2, and 3 are the sides of the triangle ABC: BC, AC, and AB, respectively. Finally, we note that in area 5 the system of conditions (2) is contradictory.

In this case, the exact solution to the problem of the upper (third) level has the form:

$$u_1^{**}=1, \quad u_2^{**}=1, \quad \rho^{**}=\frac{2}{3}$$

Let us now describe the method for solving the third-level problem for the variant of the multi-criterial model. We have:
objective functions: maximize by $x=\| x_1 x_2 x_3 \|^T \in E^3$

$$f_1(x,u)=x_1$$

$$f_2(x,u)=x_2$$

$$f_3(x,u)=x_3$$

subject to $x_1\ge 0, x_2\ge 0, x_3\ge 0,\ \ a_1(u)x_1+a_2(u)x_2+a_3(u)x_3\le b(u)$
where

$$\begin{cases} a_1(u)=u_2(3-u_1-u_2) \\ a_2(u)=u_1(3-u_1-u_2) \\ a_3(u)=u_1u_2 \\ b(u)=u_1u_2(3-u_1-u_2) \end{cases}$$

with

$$u = \| u_1 u_2 \|^T \in \Theta = \left\{ u \middle| \begin{matrix} 0.1 \le u_1 \le 2.5 \\ 0.1 \le u_2 \le 2.5 \end{matrix} \right\}$$

Let us introduce the notations:

$$\begin{gathered} y_1 = -x_1 \\ y_2 = -x_2 \\ y_3 = -x_1 \\ y_4 = a_1 x_1 + a_2 x_2 + a_3 x_3 - b \end{gathered}$$

and as a penalty function, we take $P(\tau, s) = \tau exp\left(\frac{s}{\tau}\right)$.

Then, the auxiliary functions for the one-criterion problems (8) will be:

$$A_k(\tau, x, u) = x_k - \sum_{j=1}^{4} \tau exp\left(\frac{y_j}{\tau}\right) \qquad k = 1,2,3 \tag{15}$$

The stationarity conditions of the auxiliary functions by the components of x give the following equations:

$$\frac{\partial A_k}{\partial x_j} = \delta_{kj} + exp\left(\frac{y_j}{\tau}\right) - a_j exp\left(\frac{y_4}{\tau}\right) \qquad \forall k = 1,2,3 \quad \forall j = 1,2,3$$

Finally, the derivatives of the smoothed approximating functions $\bar{F}_k(u)$ over the components of the parameter vector u are found from the relations (10)

$$\left(\bar{F}_k\right)'_{u_p} = \frac{\partial A_k}{\partial u_p}\left(\tau, \bar{x}_k(u), u\right) = \left(\bar{x}_{k1}\frac{\partial a_1}{\partial u_p} + \bar{x}_{k2}\frac{\partial a_2}{\partial u_p} + \bar{x}_{k3}\frac{\partial a_3}{\partial u_p} - \frac{\partial b}{\partial u_p}\right) exp\left(\frac{y_4}{\tau}\right) \forall k = 1,2,3 \ \forall j = 1,2 \tag{16}$$

Let us now consider the problem (11) — optimization of the mismatch of the model criteria. In our case, this problem has the following form:

minimize ρ according to the set of variables $\{x_1, x_2, x_3, \rho\}$
subject to $x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, \rho \ge 0$ and $a_1 x_1 + a_2 x_2 + a_3 x_3 \le b$ (17)

In addition, the following inequalities must be satisfied:

$$x_k \ge \bar{F}_k - \rho \qquad \forall k = 1,2,3$$

This problem can also be solved by using the method of smooth penalty functions with the same $P(\tau, s)$, for which it will be convenient to introduce (in addition to the previously defined) the notation $Y_k = \bar{F}_k - \rho - x_k \quad \forall k = 1,2,3$. In this case, the auxiliary function (12) will be:

$$E(\tau, x, \rho, u) = -\rho - \tau exp\left(\frac{-\rho}{\tau}\right) - \sum_{k=1}^{3} \tau exp\left(\frac{Y_k}{\tau}\right) - \sum_{j=1}^{4} \tau exp\left(\frac{y_j}{\tau}\right) \tag{18}$$

The conditions of stationarity for the auxiliary function (13) in (18) take the form of a system of equations:

$$\begin{cases} \dfrac{\partial E}{\partial \rho} = -1 + exp\left(\dfrac{-\rho}{\tau}\right) - \sum_{k=1}^{3} \tau\, exp\left(\dfrac{Y_k}{\tau}\right) = 0 \\ \dfrac{\partial E}{\partial x_k} = exp\left(\dfrac{Y_k}{\tau}\right) + exp\left(\dfrac{y_k}{\tau}\right) - a_k exp\left(\dfrac{y_4}{\tau}\right) = 0 \quad \forall k = 1,2,3 \end{cases} \tag{19}$$

The solution of the system (19) is denoted by $\bar{\bar{\rho}}(u)$ and $\bar{\bar{x}}(u)$, and the function below is used as the smooth approximation of the dependency $\rho^{**}(u)$:

$$\bar{\bar{E}}(u) = -E\left(\tau, \bar{\bar{\rho}}(u), \bar{\bar{x}}(u), u\right) \tag{20}$$

The derivatives of this function with respect to the components of the parameter vector u are found from formulas (14) and for the considered example have the form:

$$\left(\bar{\bar{E}}(u)\right)'_{u_p} = exp\left(\frac{y_4}{\tau}\right) \cdot \frac{\partial y_4}{\partial u_p} + \sum_{k=1}^{3} exp\left(\frac{Y_k}{\tau}\right) \frac{\partial Y_k}{\partial u_p} =$$

$$= exp\left(\frac{y_4}{\tau}\right)\left(\sum_{j=1}^{3} x_j \frac{\partial a_j}{\partial u_p} - \frac{\partial b}{\partial u_p}\right) + \sum_{k=1}^{3} exp\left(\frac{Y_k}{\tau}\right) \frac{\partial \bar{F}_k}{\partial u_p} \quad \forall p = 1,2$$

The values of the derivatives of the functions $\bar{F}_k(u)$ are found from the formulas (16).

Now, for searching for the stationary points of the function $\bar{\bar{E}}(u)$ in the space of parameters, it is possible to use any iterative method that improves the directions which can be found by using the derivatives of the first order. As an example, for the solution of problem (17) we apply standard gradient scheme of steepest ascent.

Let $t = 0,1,2,\ldots$ — iteration number. Then this scheme can be described by using the relations:

$$u_p^{t+1} = u_p^t + s^t w_p^t$$

where

$$w_p^t = \frac{1}{N_{grad}} \left(\bar{\bar{E}}(u)\right)'_{u_p} \qquad p = 1,2.$$

The value of the norm of the gradient is calculated from the usual formula of the orthonormal basis:

$$N_{grad} = \sqrt{\left(\left(\bar{\bar{E}}(u^t)\right)'_{u_1}\right)^2 + \left(\left(\bar{\bar{E}}(u^t)\right)'_{u_2}\right)^2}.$$

and the quantity s^t — the step along the improving direction for each iteration by the dichotomy method.

The results are shown in tables 1 and 2.

Remarks on the use of algorithms of second order

In conclusion, we consider the method of finding elements of the Hessian matrix for the function $\bar{\bar{E}}(u)$, the knowledge of which will allow us to use second-order methods in the search for stationary points.

Table 1

t	u_1	u_2	$\overline{\overline{E}}$	$\overline{\overline{\rho}}$
0	0.700000000	1.600000000	0.580923855	0.545812501
1	0.700000000	1.200000000	0.633041421	0.596126653
2	0.900876900	1.159092100	0.654535635	0.621144210
3	0.941239800	1.000136600	0.660271356	0.626927551
4	0.981719000	1.018621200	0.661487390	0.628152833
5	1.000234400	1.000052500	0.661620557	0.628286990
6	1.000071000	0.999946090	0.661620583	0.628287016

Table 2

t	N_{grad}	w_1	w_2	s
0	0.231363725	-1.8973 * 10^-9	-1.000000000	0.400000000
1	0.131142557	0.979887531	-0.199550560	0.205000000
2	0.076755604	0.246115279	-0.969240563	0.164000000
3	0.050581430	0.909645878	0.415384614	0.044500000
4	0.010207109	0.706089900	-0.708122202	0.026222500
5	2.43444 *10^-4	-0.837971176	-0.545714493	0.000195000
6	3.73575 *10^-5	-0.922621240	0.385707203	0.000076500

Applying the rules for differentiating a composite function to the function (20), we obtain:

$$\left(\overline{\overline{E}}(u)\right)''_{u_p u_q} = \frac{\partial^2 \overline{\overline{E}}}{\partial u_p \partial u_q} + \sum_{j=1}^{n} \frac{\partial^2 \overline{\overline{E}}}{\partial u_p \partial x_j} \frac{\partial x_j}{\partial u_q} + \frac{\partial^2 \overline{\overline{E}}}{\partial u_p \partial \rho} \frac{\partial \rho}{\partial u_q} \qquad p,q = [1,r] \tag{21}$$

The second partial derivatives are calculated directly at the point $\{\overline{\overline{x}}, \overline{\overline{\rho}}\}$, and the first derivatives, i.e. $\frac{\partial x_j}{\partial u_q}$ and $\frac{\partial \rho}{\partial u_q}$ are found according to the well-known implicit function theorem from the system of linear equations:

$$\begin{cases} \dfrac{\partial^2 E}{\partial \rho^2} \dfrac{\partial \rho}{\partial u_q} + \displaystyle\sum_{i=1}^{n} \dfrac{\partial^2 E}{\partial \rho \partial x_i} \dfrac{\partial x_i}{\partial u_q} + \dfrac{\partial^2 E}{\partial \rho \partial u_q} = 0 \\ \dfrac{\partial^2 E}{\partial x_j \partial \rho} \dfrac{\partial \rho}{\partial u_q} + \displaystyle\sum_{i=1}^{n} \dfrac{\partial^2 E}{\partial x_j \partial x_i} \dfrac{\partial x_i}{\partial u_q} + \dfrac{\partial^2 E}{\partial x_j \partial u_q} = 0 \quad \forall j = [1,n] \end{cases}$$

which is obtained by successively differentiating the stationarity conditions (13) according to the variables ρ and x_j $j = [1,n]$.

Finally, it must be mentioned that to calculate the derivatives (21) it is also necessary to know the values of the second derivatives of the functions $\overline{F}_k(u)$. These values can be found (similar to the one used above) by the method from formulas (10) and conditions (9) — the stationarity of the functions $A_k(\tau, x, u)$.

References

Fiacco, A. V., & McCormick, G. P. (1968). *Nonlinear Programming: Sequential Unconstrained Minimization, Techniques*. New York, NY: John Wiley and Sons.

Lotov, A. V., & Pospelov, I. I. (2008). *Mnogokriterial'nye zadachi prinyatiya reshenii* [Multi-criterial decision-making tasks]. Moskva, Rossiâ: MAKS Press.

Umnov, A. E. (1975). Metod shtrafnykh funktsii v zadachakh bol'shoi razmernosti [The method of penalty functions in problems of large dimension]. *Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki*, *15*(6), pp. 1451–1463.

Zhadan, V. G. (2014). *Metody optimizatsii. Chast' 1. Vvedenie v vypuklyi analiz i teoriyu optimizatsii* [Methods of optimization. Part 1. Introduction to convex analysis and optimization theory]. Moskva, Rossiâ: MFTI.

Оптимизация формы множества Парето в задачах многокритериального программирования

Була Аксел Катенди[1], Е.А. Умнов[2], А.Е. Умнов[2]

В работе рассматривается схема использования метода гладких штрафных функций для исследования зависимости решений задач многокритериальной оптимизации от параметров. Приводится описание алгоритмов, основанных на методе гладких штрафных функций, решения задачи оптимизации по параметрам уровня согласованности целевых функций и выбора соответствующей формы множества Парето.

Ключевые слова: задача многокритериального параметрического программирования; множество Парето; метод гладких штрафных функций; задача оптимизации по параметрам
JEL classification: C61, C65

[1,2] Московский физико-технический институт (Государственный университет), Москва, Россия
[1] axelbula@gmail.com, [2]mail@umnov.ru

Review of Business and Economics Studies
2018, Vol. 6, No. 1, 17–28

DOI: 10.26794/2308-944X-2018-6-1-17-28

Does Enterprise Value Really Depend on WACC and Free Cash Flow? The Evidence of Irrationality from the Oil and Gas Sector

Pavel E. Zhukov*

*Department of Corporate Finance and Corporate Management, Financial University under the Government of the Russian Federation, Moscow, Russian Federation; paul-joukov@yandex.ru

Abstract

The main objective is to check whether traditional DCF model, based on stable rational expectations for cash flows and discount rates really works in an intermediate term—from a quarter to three years. The sample was formed from six major companies of oil and gas sector. The main conclusions are—changes of enterprise value are independent of the changes WACC, free cash flow, and operating cash flows. This may be explained by the impossibility to make durable assessment neither for expected cash flow nor for the discount rate, which in fact means failure of strongly rational models like CAPM or MM. To handle out this implied irrationality the new model proposed, based on the stochastic cost of capital, which follows the model of generalized method of moments by J. Cochrane.

Keywords: enterprise value; WACC; stochastic discount rates; generalized method of moments; rationality; behavioral economics

JEL classification: C65, G32

How cash flows and cost of capital should influence enterprise value if MM theory was realistic?

The original motivation for this work is the question—how can traditional DCF approach (based on MM theory) explain real empirical middle-term changes in enterprise value? And particularly interesting is intermediate term—from a quarter to three years, which conforms to typical investment horizon for financial investments. Similar questions are discussed in the numerous investigations, but mainly for the long-term horizon and in the context of optimal capital structure, like in Bhamra, Lars-Alexander, and Strebulaev (2010) (for wider reference list see Koller, Goedhart, and Wessels (2010), or handbook by Pratt and Grabowski (2008)).

There is well known (e.g. Cochrane, 2005), that at very short-term periods (like days or weeks) price movement conforms to a random walk (or martingale). Also, there are empirically justified long-term returns on stocks and indexes (e.g. Hansen and Heaton, 2008; Chen and Hill, 2013) which are different for different periods because of business cycles and macroeconomic shocks. That may be explained by a very evident reason—growing companies should grow both in cash flows and value. And falling companies should fall in both dimensions too. However, that is not so evident for free cash flow, as growing companies (e.g., Apple) may reinvest in the growth major part of its operating cash flow and their free cash flow may not change.

And for an intermediate horizon (here it is implied from one quarter to three years) the relations between cash flows and value still is not researched well and for a good reason, because this is a puzzle. Moreover, for medium term (up to 5 years) that is a puzzle too. However, the intermediate term from one quarter to three years may conform to typical investment horizon for majority of financial investors and it may be supported by maximal samples of data, available in "Bloomberg" (e.g., for such externalities, as WACC, FCF, CFO, M-cap and etc. for the period 2000–2017).

MM theory and interrelated CAPM are widely and rightly considered as the basement of modern financial theory (see, for example, Pagano, 2005) and both have some different modern variations. The third "cornerstone" probably is Black-Sholes theory. All the three are widely applied and all three are widely questioned.

There are several possible ways to prove MM theorems. Most of them usually follow original approach as it was proposed by Modigliani and Miller (1958), and later, in a corrected form by Modigliani and Miller (1963). Commonly proofs of MM theorems are based on the impossibility of arbitration and thus assume financial markets as perfect, all-knowing and being always right. Perhaps, the most simple and elegant proof is proposed in the academic textbook (monograph in fact) by Tirole (2006) in a context with an extensive literature review on the MM theory, its latest versions, redactions and empirical verifications.

Stiglitz (1969) criticized MM theory stating five limitations of its proof, and particularly (at number five) that it was not clear how the possibility of bankruptcy affected the validity of MM theorems. Actually, MM authors and their followers did not consider the costs of bankruptcy (or financial distress) as a significant factor. So, the main direction for following MM research has become tax effects accounting—see Miller (1988). But still, many authors rightly considered the impossibility of bankruptcy (going-on concern) as a controversial. The contradiction is evident—if capital structure consists of debt only, then equity is zero and firm should go bankrupt both theoretically and in practice. Then a cost of bankruptcy was later embedded in the MM theory by "trade-off" theory as an exogenous factor.

Merton (1974) provided a specific model for the cost of bankruptcy, based on Black-Sholes theory, which still is in use by Moody's KMV model. However, Merton's work is widely (and rightly) recognized as classical and brilliant, it also consists of some contradictions and controversial approaches. For example, Merton claimed that MM theory works even under bankruptcy which is evidently wrong (see above).

The practical cornerstone of Moody's KMV model is the use of "implied" volatility instead of real volatility. And that "implied" volatility (Moody's KMV model use mainly "EDF" and "distance to default") actually is calculated back with Black-Sholes model by an empirical database which is confidential by commercial reasons. So, employing that database is the valuable secret of Moody's KMV model. The same method is applied in the Black-Sholes model for options. In the work of Zhukov (2014), there were shown, that if use real volatility instead of "implied" one there is no relation between volatility and return.

Many authors are referring to the classical monograph by Donaldson (1969), where discussed various financial strategies used by US companies and many practical examples. Donaldson concludes that companies usually stick to permanent capital structure and if change it (which is done under circumstances only), then follows the certain hierarchy of decisions. Later, his theory gets the name "pecking order theory". Following Donaldson results, Myers (1984) introduced a new direction in theory development—accounting for transaction costs. However, Myers denied the materiality of bankruptcy costs as authors of MM theory did.

Empirical check for both trade-off and pecking order theories in the view of capital structure was run by Fama and French (2002). As a result, the panel was obtained with conflicting conclusions, where some (3 conclusions) are rather in consent with the first theory, and some another (3 conclusions) are rather in consent with the second one. So the final judgment was made that both theories may be right (despite logical contradictions in the basic assumptions).

An alternative method of stochastic modeling was proposed by Strebulaev (2007). It was based on modeling a random change of enterprise value and then changing capital structure for better. That method of stochastic modeling largely is based on Merton's model for valuing bonds and default costs as an option. Strebulaev finds the contradiction between the theories of "compromise" (trade-off) and "hierarchy" (pecking order) is inconsequential in terms of modeling results. Thus, the conclusion was evolved that results of Fama and French (2001) are not controversial as it seems and really both theories may be right.

Evidently, the limited lifetime of the company may influence its WACC and the enterprise value compared with a firm unlimited in time, as it supposed to be under the "going-concern" concept. That was an initial idea for the theory of Brusov-Filatova-Orechova, proposed in Brusov et al. (2011). It was based on optimizing the enterprise value with a limited lifetime, versus a model firm in the MM theory which was denoted as "perpetuity" firm.

An alternative view of the corporate financial policy was exposed by Tirole (2006), where the central point is an agency conflict and related information asymmetry between insiders (managers) and outsiders (shareholders). That is certainly one of the main sources of market imperfectness and with no doubt one of the central (if not the main) problem for the corporate finance. The agency conflict obviously may have an impact on both costs of bankruptcy (or financial distress) and transaction costs. The monograph is based on the wide review of empirical results and considers some contradictions between empirical results and financial theory. That goes well beyond of either MM or CAPM theories, even modernized.

The common identity for the enterprise value (e.g., see Damodaran (2008)) is:

$$EV = MV(Eq) + MV(ND) \quad (1)$$

The second term represents the market value of the net debt and usually, it is the difference between gross debt and liquid assets — cash and market securities.

$$MV(ND) = MV(D) - Cash - MS \quad (2)$$

From the postulate of investor's rationality there can be deduced that enterprise value is equal to the discounted value of the free cash flows to the company:

$$EV(t0) = \sum_{t=t0}^{\infty} FCF(t)/(1 + CC(t))^{t-t0} \quad (3)$$

Here FCF(t) — the expected future free cash flows to the firm, and CC(t) is the expected cost of capital. EV(t0) is the enterprise value at the moment t0.

The company's cash flow (free cash flow) is usually defined as the cash available for distribution to investors and creditors after capital expenditures:

$$FCF(t) = CFO(t) - CAPEX(t) + Int(t)\,(1-T(t))$$

Here CFO(t) is net operating cash flow; CAPEX(t) — net investment in fixed capital; Int(t) — interest for the loan; T(t) — the effective tax rate applicable.

Identity (3) is also widely used as justification for the company's valuation methods on discounted free cash flows or DCF (e.g., see monograph by Koller, Goedhart, and Wessels, 2010). Accordingly, both in theory and in practice, the central problem of the company value management is usually reduced to cash flow management, risk management or capital structure management.

Formally (3) may also be treated as identity, or as an equation for either unknown cash flow or capital cost. However, as there is only one equation, one may find either single average cash flow for the given cost of capital or single average capital cost for given average cash flow or given both find enterprise value and etc.

MM theory in fact (albeit implicitly) use the postulate of a rational behavior of investor which create enterprise value (3) by the market equilibrium price. And it proves that the discount rate in (3) is equal to the weighted average cost of capital WACC, composed from required return (opportunity costs) on shareholders' equity and the required yield on debt, taken after tax shields:

$$WACC(t) = Re(t) \times we + Rd(t) \times wd(1 - T) \quad (4)$$

Here Re — required return for equity (commonly treated as a return to a diversified portfolio with the same risk and leverage); Rd — the cost of interest-bearing long-term debt; *we* and *wd* — shares of equity and debt in the enterprise value; T — marginal rate of corporate income tax.

As a rule of thumb, the weighted average cost of capital (4) is considered permanent in the MM theory since all variable factors (including individual risks) are counted in the expected free cash flows.

Expression (4) for the discount rate, and especially combined with CAPM for the cost of equity may be considered as the very arguable part of MM theory. But in fact, the expression (3) is arguable too.

However (3) may be considered just as an identity for enterprise value, as it directly follows from the postulate of the rational behavior. Also, identity (3) can be used as the equation for average discount rate given the cash flows or as the equation for average cash flow at a specified discount rate and growth rate. In the work by Zhukov (2015), there was presented another alternative and generalized variant for MM theory, including all bankruptcy (default, financial distress) costs, and transaction costs as adds to discount rates, covering both trade-off and pecking order theories. There is shown that modified MM theory can be built just on identities (3) and (4), added with an assumption that required yield on equity depends linearly on the debt leverage. The latter follows directly from the effect of financial leverage. So, in fact, MM theory actually relies neither on the impossibility of arbitrage, nor on "going concern", as it is widely accepted, but just on (3) and (4) identities, which may be derived almost directly from the postulate of rational behavior for investors.

The important generality of (3) is—free cash flow is a stochastic process and may depend on the time (t). For the median value there may be considered any of a reasonable trend (linear, exponential, declining and etc.). And the stochastic error may have any distribution but must have zero median. But the central question for the (3) is—may either cash flows or discount rates depend on the reference point (t0)?

This is important because investors' expectations regarding cash flows and risks may change. Formally in the (3) and (4), this independence is not required, as the model is designed just for an assessment at the current moment (t0) and every time vision of future may change. But there are two central points—how that change may be reflected in a model and how it may affect outcomes?[1] Formally if μ is probability measure and $\lambda(t, t0, \mu)$ is probability density distribution for FCF, then:

$$\mathrm{FCF}(t, to) = \int_{-\infty}^{\infty} \lambda(t, t0, \mu)\, d\mu$$

However, if the mathematical expectation of stochastic cash flows FCF(t, t0) may change with the reference point of time (*t0*) then this parametric set of stochastic processes (*t* is a parameter) is not stationary with time (t0). And then the cash flow as the stochastic process cannot be represented as a sum of a trend functions of future moment *t* and any stochastic process (may depend on *t0*) with zero median. Which means—FCF(t, t0) may form stochastic process with unknown and changing variance, unknown and changing median and also with unknown and changing trend.

All that means that investor's appraisal at any time *t0* is not durable (prudent) and as unreliable as the bet in the casino. But if transaction costs in casino come lower than in the stock market (due to lower costs of expert services), the stock market would go bankrupt. So far, as the stock market still survive the competition with a casino, it means that either it provides lower transaction costs or that it provides some other service besides casino do. For that reason, if rationality of investors prevails over irrational (changing) assessments, expected free cash flows and discount rates in (3) must be independent on reference point *t0*—the moment of appraisal. And this important statement forms the concept denoted here as "strong rationality".

Definition 1. Strong rational model (for a strong rational investor) for the enterprise value is any model, based on identity (3), where mathematical expectations for free cash flows to the firm (stochastic) and capital cost (fixed) both may depend on the time, but do not change with relevance point t0.

So strong rationality assume that investor does not change its assessment (forecasting) of cash flows and discount rates through the time. The strong rational model means that investor may reliably assess future expected cash flows and then this assessment should not change with time.

The postulate of the strong rational behavior was widely criticized by behaviorists (see, e.g., Richard Thaler, 2016), but as Tirole (2006) sentenced — there is no rigorous model created to include irrationalities (except maybe Beta-Delta discounting). By the other hand, as R. Thaler stated (in the cited above book), many academicians of traditional school believes, that gesture by "invisible hand" of the market may make those rational models work. Also they evidently believe, that market price is "always right" and "fair", and market itself is "perfect", "all-knowing", and finally benevolent, while unpredictable and incomprehensible[2]. So far, applicability of the strong rational model (3), (4) for the empirical data may be the good test for rational behavior postulate.

[1] The author believes that the same questions were probably stumbling-block for the Merton (1974) work on the bond price. But probably that ideas was initially for the Merton's (1972) model ICAPM which is widely underestimated because experts in finances use multifactor models (see Maio, 2012) while academicians prefer "rational" CCAPM.

[2] These points may pose a good reason to build a cathedral for Holly Market in Chicago University where anybody may pray to the Saint Greed for Benevolent Fair Price. However one must differentiate science and religion.

On the other hand, it may provide a good starting point for the construction of alternative models if necessary.

One of the possible alternatives is to use stochastic discount rate instead of stochastic cash flows in (3). Cochrane (2011) proved that the main role in the prices volatility plays the volatility of discount rates but not the volatility of cash flows. However, those results are very general as the study was conducted for stock exchange indexes, and over a long period of time. So their applicability to individual companies and especially in a medium term is not clear.

Two methods employed—fixed discount rates with stochastic cash flows and fixed cash flows with stochastic discount rates

At the first step the strongly rational model (3) and (4), related to the MM theory (either common or generalized) was examined for applicability to empirical data. Specifically, panel research applied to find the dependence on medium-term changes of enterprise value and capitalization on medium-term changes of cash flows and WACC. There were examined relative changes which are TS-type (trend stationary) and therefore are subject to the usual F-statistics (e.g., see Hamilton, 1994; Wooldridge, 2002). And the results were certainly negative—"strongly rational" model, which implies future stochastic cash flows with fixed expected value and fixed discount rates fails.

At the second step, there was applied author's modification of the generalized method of moments proposed by Cochrane (2005). The term "generalized method of moments" and its idea comes from similar in form (but different in a purpose) general method for statistical evaluation of the best parameters for econometric models, proposed by Hansen (1982). Cochrane's generalized method of moments originally comes from CAPM "family" of methods, which is heavily based on the long-term rationality of investors. And inside the "family" (which includes CAPM, ICAPM, CCAPM and etc.), Cochrane used mostly CCAPM, which is the most "rational" as it is based on the Arrow-Debreu model for global economic equilibrium which implies very restrictive assumptions. Tirole (2006) asserts that the entire MM theory may be obtained from the same Arrow-Debreu model as well[3]. So all the modern financial theory may be evolved from the global economic equilibrium.

Cochrane (2005) used CCAPM for theoretical excerpts, while specified that theoretical approach leads to unrealistic estimates of internal parameters and CCAMP is poorly applicable to practice. And there are stated some contradictions of CCAPM outcomes with practice, which Cochrane call "puzzles", while actually, the greater puzzle would be if there were found evidence for practicability of CCAPM[4].

And (citing Cochrane) if CCAPM was really applicable to the practical economy, then economic theory may be considered accomplished, particularly because any market price for asset would be assessable with CCAPM (or CAPM, ICAPM, and etc.). Apparently, it is not.

But Cochrane states particularly that there is central equation (5) for that model which does not depend on any highly limiting assumptions (like market equilibrium, stationarity, normal distribution, the impossibility of arbitrage and etc.).

$$p = E(mx) \tag{5}$$

Here p is expected price and on the right side, there is a mathematical expectation of the product of factor vector (m) and vector of expected future return (x).

So, actually, the model (5) is not affiliated with CCAPM or CAPM or any other theory and it may be taken as the initial model itself. However use of stochastic discount factor "m" in (5) seems quite logical, as it describes a choice between the future and the present consumption.

Generally, in (5) one may choose between two alternative approaches—either to use stochastic cash flows "x" but fixed discount rates "m" (like in MM), or to run with fixed expected cash flows "x" and stochastic discount factors "m". The reason for this duality is evident—any risk factors can be taken into account either in cash flows

[3] This is not really surprising if apply mathematical logic. By the Gödel incompleteness theorems in a controversial system, any statement may be proved as true, as well as an opposite one.

[4] There is nothing impossible. As Hansen (2017) explained in his latest brilliant book even wrong theory may be justified if use optimization method for internal parameters of the model on a fixed sample. However, does it eliminate the difference between right and wrong theory? Gödel theorems claim it doesn't.

or in discount rates and both approaches should theoretically lead to the same results. Underlying for the first approach is the idea that investor (rational or not) at the time of the evaluation may not be willing to use any future required rates of return which are impossible to predict. Instead, investor rather tends to apply the current required rate of return at the time of assessment. In this case expected cash flows "x" must be considered as stochastic, while discount factors "m" are fixed. If mathematical expectation for the all future cash flows are reliably assessed by investor and expected cash flows and discount rate will not change in future, this model must be denoted as strong rational (by the Definition 1).

All the theories, that assume impossibility of arbitrage actually relies on that case of strong rationality. Particularly, so do MM, CAPM and Black-Sholes theories which constitute the foundation of the modern finances. The same is CCAPM because future consumption becomes unknown (as it really is).

For example, MM theory does not formally require (3) and (4), but it relies on the assumption that risk and expected cash flows for all the companies form the enterprise value and may be assessed for the infinite periods. If this assessment will change with the time than enterprise value becomes unreliable (or irrational) and then so becomes all the MM theory. Formally one may introduce parameter "t0" and assume all that estimates depending on time. But it would mean that market is not in equilibrium and if so, arbitrage is possible and then MM, CAPM (for CAPM case see Fama and French (2006)), and Black-Sholes theories fail. Only ICAPM by Merton (1972) and factor models may still work if they do not require equilibrium.

Some cases of weaker rationality may be captured by behavioral economics. However, Tirole (2006) stated that this new area of economics does not have proper models yet. The well-known behavioral models include hyperbolic discounting, and particularly Beta-Delta discounting model. These models certainly do not conform to the strong rational case even if beta and delta are permanent.

Generally, if either expected cash flows or discount rates in the model may change along with the time of assessment, then the model probably may describe some kind of irrationality. Effectively it means the very simple point—investors are unable to predict neither expected cash flows nor risks for the long future. This point certainly is reasonable and rational from the practical approach, but in traditional theory, it may look like irrationality.

For the second approach, the underlying idea is that investor anticipates average expected cash flows "x" as fixed (determined), possibly with permanent growth rate (positive, zero or negative). For that approach discount rates "m" must be stochastic, which relates to the generalized method of moments by Cochrane (2005) in the form (5). This case may be denoted as "weak rational case" as it implies some of the rationality, but not in the strong form.

Results for the first approach—checking common DCF model (generalized MM theory)

First, represent actual cash flow FCF*(t), observed at the time (t) as the sum of expected cash flow FCF(t) and stochastic fluctuation δ (t) with zero median[5]:

$$FCF^*(t) = FCF(t) + \delta(t) \tag{6}$$

No assumptions about distribution are made here but median must be equal to zero.

If a stochastic series (3) converges in the sense of mathematical expectation, then its sum is equal to the expected enterprise value (e.g. see Wooldridge, 2002).

Under the assumption that free cash flow model (3), (4) is the underlying basis for enterprise value, one get the equation:

$$EV(t) = (FCF(t+1) + EV(t+1))/(1 + WACC(t))$$

This also can be written in incremental form:

$$EV(t+1) - EV(t)\,(1 + WACC(t)) = -FCF(t+1) \tag{7}$$

This equation (7) may be extended from the 1 period of time (e.g. quarter or year) to any number of periods.

$$EV\,(t+1) - EV(t)\,(1 + WACC(t))^{-n} = \sum_{\tau=1}^{n} FCF\left(t+\tau\right)/(1 + WACC(t+\tau))^{\tau} \tag{8}$$

[5] No assumptions about distribution is made here but median must be equal to zero.

For the strong rational case (3) and (4), expressions (7) and (8) represent the increase in the enterprise value for the certain period. Further (8) will be considered as the model for increments in enterprise value from 1 quarter to 3 years. Following questions are examined for the selected sample of oil and gas sector:

1. Do the fluctuations in the enterprise value really may be explained with the generalized MM model (3), (4)?
2. Coming away from MM theory—are changes in enterprise value related to the changes in free cash flow, operational cash flow or WACC?
3. If change free cash flow to the operating cash flow plus interest minus tax shields, would it change the answers to questions 1 and 2?
4. To what extent changes in the market capitalization of the company may be explained by the changes in its enterprise value?

For the strong rational case (8) become deterministic. In that case, one may observe that correlation of actual enterprise values and those derived from (8) as equal to 1 and with the significance level (probability of the hypothesis H0) close to zero. However, if that correlation is equal to zero (hypothesis H0), one must assume that investors essentially change their assessment of future risks and (or) cash flows for the every projection period. So, that rational case is not the real one.

The sample for the study was compounded from six companies of oil and gas sector—Lukoil, Rosneft, Gazprom, Novatek, BP, Dutch-Shell. But in order to identify the possible impact of a sample on the results, the company from the opposite sector (by systemic risks) was added—Coca-Cola[6].

To answer the question 1 the validity of the model (8) was examined to the increment in the enterprise value after 1 quarter and up to the 3 years. The result was a conclusion that enterprise value increment for the periods from 1 quarter to 3 years with probability from 35% to 86%, is not correlated with the theoretically expected. So, H0 hypothesis of zero correlation is most likely (or, at least, can't be rejected), and rational case can't explain the real deviation of the price.

$$\mathrm{Corr}\left(EV_{(t+n)}/(1 + WACC(t + n))^{n} - EV_{t}),\ \sum_{\tau=1}^{n} FCF(t+\tau)/(1 + WACC(t + \tau))^{\tau}\right) \sim 0$$

It means also that the common MM model (3), (4) with permanent parameters does not work at the middle-term forecasting period. So, if rational investors use model (3), (4) for forecasting period, they permanently (at least every quarter) change estimates for either the future expected cash flows, or the discount rates, or both together. And these re-evaluations as stochastic process do not make any predictable trend (median), distribution or variation which makes it not stationary with any predictable trend. That is not really surprising because otherwise, the market price was more predictable than it really is. But that does not follow from the efficiency of the market because stochastic changes of cash flows may be considered as a reason for re-estimation of share prices. Eventually, they are not. The re-evaluations of future risk and returns are responsible for deviation of share prices. This is the outcome.

Question 2 is actually a generalization of question 1, but independent on the model (3), (4). Obviously, the cash flows reflect the benefits to investors. So far, under rational behavior postulate, even if investors change their expectations for cash flows and risks, they still should adjust it with any material change in cash flows, or risks, reflected in WACC. And since the cash flows of a company usually can be anticipated from the year ahead-based financial planning, this approach must be based on planned (forecasted) numbers for the future.

Also, there are numerous empirical studies confirming the intuitively obvious assumption that investors adjust prices depending on the news. And news usually relates to either the future cash flows to the company or to the macroeconomic risks. On the other hand, WACC reflects the systematic (common) risks of the company or at least pretend to do it. Accordingly, the change in the weighted average price of capital should lead to a change in the company's value not only for the model (3), (4) but also on the general basis of the rational expectations hypothesis.

But the answer to question 2 is negative—no dependence was observed for the selected sample. It can be assumed that this is true for the many

[6] Actually, the sample included more companies from 4 industries, but for the brevity this work describes results for oil and gas sector, adding just one company outside it—Coca-Cola, for comparison.

other companies. A brief description of the results of the study is added in Annex 1.

The relative increment of the enterprise value (in % to the previous) was chosen as the dependent (explained) variable, and as independent (explaining) variables there was chosen relative increments of free cash flow, net cash flows, or WACC. Note that since relative changes were chosen, this process must be TS-type[7] (trend stationary) with zero trend (e.g., Hamilton, 1994; Wooldridge, 2002). Correlation values obtained during regression range from 0.13 to 0.01, but in any case the hypothesis H0 could not be rebutted with a minimally acceptable level of significance (10%), and for the most cases, its probability is higher than 30%.

The answer to the question 3 for oil and gas sector is negative too—results for free cash flow and net cash flow from operating activities nearly coincide. This appears to be related to the relatively stable investment cash flows and is caused by the relative stability of investments in the oil and gas sector which is not the case for fast-growing companies (e.g., for Apple operating cash flow was fast increasing from 2000 to 2017 while free cash flow did not change so much).

However, answer to question 4 is positive—it was found that enterprise value and market capitalization has a strong interdependence for the oil and gas sector, which corresponds to the MM outcomes. Presumably, they are co-integrated, unless a radical change in the capital structure or risks happens. But for the company from beverages sector (Coca-Cola) that result is negative. Probably this difference is caused by the stability of capital structure in oil and gas sector.

These results are surprising and even paradoxical since it is generally assumed (see, e.g., Koller, Goedhart & Wessels, 2010), that investors adjust their estimates of enterprise value either for a change in the cash flows of the company or in the risks. In addition, since the expression for WACC (4) presumably reflects systematic risk, it turns out that the main role for the variability of enterprise value plays idiosyncratic (individual) risks, to the contrary of CAPM or CCAPM theory (e.g., see Sharpe, Alexander, & Bailey, 1999).

[7] Trend stationarity assumes zero median and fixed variation of error but does not require normal distribution, so it is often considered as a weak case of stationarity.

Specifically, consider the result that change of enterprise value is independent on those of cash flows or WACC. That looks like a puzzle itself because cash flows should reflect expected a return and WACC should reflect risk, or at least, its major part. The most likely answer may be—medium term fluctuations of cash flows and WACC are stochastic and therefore ignored by the market. If so, then—where the "holly" market get data for its "all-knowing" appraisal? The logical answer may be—investors are not as strongly rational as it assumed in MM model or in the model (3) and (4). Then, they must be using somehow another model to price assets.

Does it mean "irrationality"? Perhaps, but in a very specific way — investors can't forecast future expected cash flows and probably change their expectations or discount rates at least every quarter (year) or maybe even faster. So, if one treats "rationality" as the ability of an investor to forecast future expected cash flows and apply to those fixed (but maybe different for every period) discount rates, then answer is "Yes", it means the absence of rationality in that strong definition (see section 2).

The second step—implied stochastic discount rates

At the second step, there will be considered a model of weak rationality (5), based on stochastic discount factors and fixed cash flows.

For the start, consider the most irrational case of model (3), where all variables change every moment, and then depends on the reference point. Particularly, if (3) depends on t0, then the most irrational model may be represented in a form:

$$EV(t0) = \sum_{t=t0}^{\infty} FCF\left(t,\ t0\right)/(1 + CC(t, t0))^{t-t0}. \qquad (9)$$

Here all approximations for the cash flows and for the cost of capital may depend on the reference point, so all of them may change every moment. And, for all t0 greater, than t, discount factor CC(t, t0) may change, which means that model (9) is more general than Beta-Delta discounting model. However, this generality may not be necessary, as there is just one explained variable on the left side.

Therefore the most interesting for practical use case may just use one expected (fixed) cash flow, divided by stochastic rate:

$$EV(t0) = FCF(t0)/r(t0).$$

And stochastic rate r(t0) may be represented as a difference of stochastic implied a cost of capital CC(to) and the permanent growth rate of cash flows:

$$r(t0) = ICC(t0) – g.$$

As it was proved in the previous section, changes of enterprise value are very likely being independent of free (or operational) cash flows and WACC. So it is logical (rational) to use expected free cash flow as determined, but, maybe growing (or declining) along with any of specific trend (linear or exponential or any other). Then stochastic discount rates absorb all the information of price changes. That assumption may be derived from two following hypotheses:

1. Investor ignores random fluctuations in cash flows and instead uses some pre-determined value of the expected free cash flow.
2. The investor uses stochastic (randomly changing) discount rates reflecting either the stochastic risks or changes in the investor's expectations about the growth rate of cash flows in future.

With some basic (minimum) investment, the expected free cash flows of the company must have permanent expected value (as all the risks are considered in the discount rates), but with additional investments, free cash flows may grow. Denote FCFexp expected basic free cash flow, independent of time:

$$FCFexp = E(FCF(t)).$$

In the case of zero growth discount rate R may be found from equation:

$$E(EV(t)) = EVexp = FCFexp/R(t).$$

But with some additional investments average free cash flows (but not necessarily actual cash flows) will grow with changing growth rate g(t):

$$EVexp(t) = FCFexp(t+1)/(R(t) – g(t)). \quad (10)$$

Assumption 1 means that investors may use some pre-determined value for expected free cash flow, but dependent on time. And for every period investor may change required rate of return and (or) growth rate. Then discount rate is stochastic, although investor at every moment applies a single rate for future cash flows.

Then, for (10) with changing growth rate the stochastic discount rates are:

$$r(t) = R(t) – g(t) = FCF(t + 1)/EV(t). \quad (11)$$

Here R(t) — the stochastic cost of capital, reflecting stochastic risks, and *g* is a changing growth rate for expected free cash flow.

More generally, whatever the assessment methods are actually used by investors, their results can be summarized as the expected cash flow (FCF(t+1)) and a stochastic discount rate (r(t)). If the cost of capital was equal to WACC, then R(t) from (11) would be equal to WACC as well.

However, the results of the research show that growth rate, derived from WACC, have no relation to the changes in the enterprise value, which makes its use pointless. Moreover, the mean values of WACC are much higher than empirical stochastic discount rates (see table 1) and that difference can't be explained by any growth rate (which is negative for the oil sector). While it is interesting that the standard deviation of stochastic rates still is roughly equal to that of WACC. The same results are shown by other peer companies.

So judging by the empirical data, the optimal option for practical appraisal is to use stochastic discount rates. Explaining its meaning both cost of capital and expected growth rate may be chosen as stochastic variable as well as. Both of them may reflect an assessment of risks and prospects by investors and investors may either assess both of it or just use one stochastic discount rate instead. However, the cost of capital looks like better explainable and logical variable for risk assessment, which may change every period on the basis of new data. And if the cost of capital is chosen as a stochastic variable the change of expected growth rate is an unnecessary complication. Therefore growth rate may be considered as permanent without loss of generality.

Conclusions

1. Model (3), (4) (a generalization of MM theory) is equivalent to the strong rational investor, so it is a strong rational model (see definition 1).
2. Strong rational model (3), (4) does not explain the medium-term changes of enterprise value.

Table 1
Averages and variability of WACC and stochastic discount rates for BP since 2000 by 2016. (according to data from Bloomberg)

	WACC	CFO, mln. $	FCF, mln. $	Rcfo	Rfcf	EV, mln. $	Mcap, mln. $
Median	0.088	539	135	0.012	0.003	144000	111000
St. Var.	0.190	1.840	0.460	0.210	0.210	0.270	0.380

Note. Rcfo — stochastic discount rate (R(t)) derived from CFO; Rfcf — the same rate, applicable to FCF. As free cash flow is four times smaller its discount rate is four times lower.

3. In the medium-term assessment using the WACC calculated in line with the theories of MM and CAPM (or CCAPM), as a discount rate for the free cash flow gives no results comparable with empirical data for a selected sample of the companies.
4. Changes of the WACC do not affect medium-term changes in the enterprise value or the capitalization for the selected sample of companies which means that WACC does not reflect the real cost of capital for the selected sample.
5. As changes of WACC reflect changes in systematic (common) risk, it may be assumed that individual (idiosyncratic) risk (which are not reflected in WACC) provide a major influence on medium-term changes of enterprise value.
6. Medium-term changes in free cash flows do not affect changes to the enterprise value of the companies from the selected sample (which deliberately does not include growing companies).
7. Cochrane's model of generalized moments in the form (5) may be useful to evaluate a company price instead of models with fixed discount rates and stochastic cash flows, like (3), (4).
8. To assess the enterprise value of company the possible way is to use long-term fixed cash flows, growing with permanent growth rate (positive, zero or negative) and stochastic cost of capital.

Appendix

Checking the independence of price changes from the change of the company's cash flow (FCF, CFO) and the discount rate (WACC)

Because the company's value changes constantly in real time, it can be interpreted as a realization of a stochastic process. Changes in the prices of shares in companies usually are of type DS (difference stationary) and to examine them usually there are used autoregressive models (AR), or combined with MA (moving average) process—ARIMA models. In the present work subject of research is the dependence of relative change of enterprise value from the relative change in cash flows and WACC. The percentage change in asset prices refers to processes of type TS (trend stationary) with zero trend and, therefore, results may be assessed with applicable F-statistics. For the percentage change in the enterprise value it is:

$$dEV(t) = (EV(t) - EV(t-1))/EV(t).$$

Independent variables were changes the discount rate WACC, free cash flow FCF (4) and operating cash flows adjusted for interest payments (10):

$$dWACC(t) = (WACC(t) - WACC(t-1))/WACC(t)$$
$$dCFO(t) = ((CFO\ (t) - CFO(t-1))/CFO(t)$$
$$dFCF(t) = ((FCF(t) - FCF(t-1))/FCF(t).$$

For example, for the BP Corporation, the chance of hypothesis H0 is over 84%. The only variable which tends to show a sustained and significant correlation with the enterprise value (and with a correlation coefficient close to one) is market capitalization. This conclusion is consistent with MM. However, this conclusion is not trivial, given that in expressions (2) and (3) all variables may change significantly over time. Moreover, for the one company from the sample, Coca-Cola, this conclusion turned out to incorrect—change in enterprise value was not associated with changes in capitalization. The reason for this is not clear, but it is clear that this company proved an exception to the general rule (probably due to the nature of its financial policy).

Also for this company, there are significant dependence of the change in market capitalization and free cash flow changes (for other companies it is not).

Thus, there are no observed significant dependencies of changes in enterprise value from changes to FCF, CFO, and WACC. On the contrary, it is very likely that estimated correlation coefficient is indistinguishable from zero.

Table 2
According full price company (probability of the hypothesis H0 and R 2)

Company	FCF (p-val.)	CFO (p-val.)	WACC (p-val.)	R^2	F-stat (p-val.)	MCAP (p-val.)	R^2 for the Mcap
BP	0,64	0.60	0.78	0.01	0.89	10E-57	0.97
Shell	0.35	0.38	0.50	0.13	0.07	-331.57	0.94
Coca-Cola	0.61	0.65	0.95	0.01	0.97	0.98	1.6E-05
Rosneft	0.63	0.32	0.14	0.07	0,40	2,6E-27	0.94
Lukoil	0.31	0.71	0,40	0.02	0.68	5,1E-45	0.96
Gazprom	0.85	0.24	0.38	0.07	0.23	2,62E-27	0.94

References

Bhamra, H., Lars-Alexander, K., & Strebulaev, I. (2010). The aggregate dynamics of capital structure and macroeconomic risk. *Review of Financial Studies, 23*(12), pp. 4187–4241.

Brusov, P., Filatova, T., Orehova, N., & Brusova, N. (2011). Weighted average cost of capital in the theory of the Modigliani-Miller, modified for a finite lifetime company. *Applied Financial Economics, 21*(11), pp. 815–824.

Cochrane, J. H. (2011). Presidential Address: Discount Rates. *The Journal of Finance*, *4,* pp. 1047–1108.

Cochrane, J. H. (2005). *Asset pricing*. Princeton, NJ: Princeton University Press.

Cochrane J.H. (2014). A Mean-Variance Benchmark for Intertemporal Portfolio Theory. *The Journal of Finance*, Vol. LXIX, No. 1, pp. 1–49. Retrieved from https://faculty.chicagobooth.edu/john.cochrane/research/papers/jofi12099.pdf.

Damodaran, A. (2008). *Investment Valuation*. New York, NY: John Wiley and Sons.

Chen, J., & Hill, P. (2013). The impact of diverse measures of default risk on UK stock returns. *Journal of Banking & Finance*, *7*, pp. 5118–5131.

Davydenko, S., Strebulaev, I., & Zhao, X. (2012). A Market-Based Study of the Cost of Default. *The Review of Financial Studies*, *25*(10), pp. 1024–1033.

Donaldson, G. (1969). *Strategy for Financial Mobility*. Cambridge, MA: Division of Research, Harvard Graduate School of Business Administration, Harvard University.

Denzler, S.M., Dacorogna, M.M., Müller, U.A., & McNeil, A.J. (2006). From default probabilities to credit spreads: Credit risk models do explain the market price. *Finance Research Letters, 3*, pp. 79–85.

Fama, E., French, K. (2002). Testing tradeoff and pecking order predictions about dividends and debt. *Review of Financial Studies*, *15*, pp. 1–33.

Fama, E., & French, K. (2006). The Value Premium and the CAPM. *The Journal of Finance*, *5*, pp. 2163–2185.

Frank, M., & Goyal, V. (2008). Trade-off and pecking order theories of debt. In B. Espen Eckbo (Ed.), *Handbook of Empirical Corporate Finance* (Chapter 12, pp. 135–197). Vol. 2. Amsterdam, the Netherlands: Elsevier/North Holland.

Froot, K., & O'Connell, P. (2008). On the pricing of intermediated risks: Theory and application to catastrophe reinsurance. *Journal of Banking and Finance, 32*, pp. 69–85.

Garleanu, N., & Pedersen, L. (2011). Margin-based asset pricing and deviations from the law of one price. *Review of Financial Studies, 24*, pp. 1980–2022.

Graham, J., & Leary, M. (2011). A review of empirical capital structure research and directions for the future. *Annual Review of Financial Economics*, *3*, pp. 309–314.

Hansen, L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, *50*(4), pp. 1029–1054.

Hansen, L., Heaton, J., & Li, Nan. (2008). Consumption Strikes Back? Measuring Long-Run Risk. *Journal of Political Economy*, *116*(2), pp. 260–302.

Hansen, L., & Sargent, T. (2016). *Robustness*. Princeton, NJ: Princeton University Press.

Hamilton, J. (1994). *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Koller, T., Goedhart, M., & Wessels, D. (2010). *Valuation: Measuring and Managing the Value of Companies*. Hoboken, NJ: John Wiley & Sons, Inc.

Maio, P. (2012). Multifactor models and their consistency with the ICAPM. *Journal of Financial Economics, 106*, pp. 586–613.

Merton, Robert C. (1973). An intertemporal Capital Asset Pricing Model. *Econometrica*, *5*, pp. 867–887.

Merton, Robert C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance, 29*(2), pp. 449–470.

Miller, M. (1988). The Modigliani-Miller Proposition after Thirty Years of Economic Perspectives. *Journal of Economic Perspectives*, *2*, pp. 99–120.

Modigliani, F, & Miller, M. (1958). The cost of capital, corporation finance and the theory of investment. *American Economic Review*, *48*(3), pp. 261–297.

Modigliani, F., & Miller, M. (1963). Corporate Income Taxes and the Cost of Capital: A Correction. *American Economic Review*, *53*, pp. 147–175.

Moody's KMV model. (n.d.). Retrieved from www.defaultrisk.com.

Myers, S. (1984). The capital structure puzzle. *Journal of Finance*, *39*, pp. 575–592.

Myers, S. (2001). Capital Structure. *Journal of Economic Perspectives*, *15*, pp. 81–102.

Pagano, M. (2005). The Modigliani-Miller Theorems: A Cornerstone of Finance. University of Salerno, *Working Paper No. 139*.

Pratt, S., & Grabowski, R. (2008). *Cost of capital: applications and examples* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Strebulaev, I. (2007). Do tests of capital structure theory mean what they say? *The Journal of Finance*, *62*, pp. 1747–1787.

Stiglitz, J. E. (1969). A Re-Examination of the Modigliani-Miller Theorem. *American Economic Review*, *59*, pp. 784–793.

Sharpe, W. F., Alexander, G. J., & Bailey, J. V. (1999). *Investments* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Tirole, J. (2006). *The theory of corporate finance*. Princeton, NJ: Princeton University Press.

Thaler, Richard H. (2016). *Misbehaving. The making of behavioral economics*. New York, NY: W.W. Norton & Company.

Wooldridge, M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). Cambridge, MA: The MIT Press.

Zhukov, P. (2015). Modification of the theory of capital structure, taking into account the probability of bankruptcy. *Financial Analytics: Problems and Solutions*, *269*(35), pp. 50–60.

Zhukov, P. (2014). Default risk and its effect on a bond required yield and volatility. *Review of Business and Economics Studies, 4*, pp. 87–98.

Разве стоимость компании действительно зависит от средневзвешенной стоимости капитала и свободного денежного потока? Свидетельства иррациональности в нефтегазовом секторе

Павел Е. Жуков[1]

Основная задача статьи — проверка, действительно ли традиционная модель DCF (Discounted Cash Flow — модель дисконтированных денежных потоков), основанная на стабильных рациональных ожиданиях денежных потоков и ставок дисконтирования, работает в промежуточном периоде — от квартала до трех лет. Выборка была сформирована из шести крупных компаний нефтегазового сектора. Основные выводы — изменения стоимости предприятия не зависят от изменения средневзвешенной стоимости капитала, свободного денежного потока и текущего денежного потока. Это может быть объяснено невозможностью осуществления долговременной оценки ни ожидаемых денежных потоков, ни учетной ставки, что фактически означает несостоятельность строго рациональных моделей CAPM (Capital Asset Pricing Model — Модель оценки долгосрочных активов) или ММ (Market Model). В статье предлагается учесть подразумеваемую иррациональность в новой модели, базирующейся на стохастической стоимости капитала, связанной с моделью обобщенного метода моментов J. Cochrane (ОММ; англ. GMM — Generalized Method of Moments).

Ключевые слова: стоимость предприятия; ССП; стохастические ставки дисконтирования; обобщенный метод моментов; рациональность; поведенческая экономика

JEL classification: C65, G32

[1] Департамент корпоративного финансирования и корпоративного управления, Финансовый университет, Москва, Россия; paul-joukov@yandex.ru

Review of Business and Economics Studies
2018, Vol. 6, No. 1, 29–42
DOI: 10.26794/2308-944X-2018-6-1-29-42

Weather Derivatives in Russia: Farmers' Insurance against Temperature Fluctuations

Eric Carkin[2], Stanislav Chekirov[3], Anastasia Echimova[3], Caroline Johnston[1,2], Congshan Li[2], Vladislav Secrieru[3], Alyona Strelnikova[3], Marshall Trier[2], Vladislav Trubnikov[3]

[1]main draft editor, derivative pricing research
[2]Worcester Polytechnic Institute, MA, USA
[3]Financial University under the Government of the Russian Federation, Moscow, Russia

Abstract
This project proposes the use of weather derivatives, a type of financial instrument with a payout based on weather conditions, as a method for Russian farmers to hedge against daily temperature fluctuations. We created a weather derivative simulation tool in Microsoft Excel that calculates the effect of temperature on crop yield and then analyzes how the return of weather derivatives can potentially compensate for crop loss. Based on this tool, we developed a series of recommendations to help implement this system of protection with real users.
Keywords: weather derivatives; temperature fluctuations; hedge; crop loss
JEL classification: G13, G17, Q18

Utilizing Weather Derivatives in Russia

In 1998 it was estimated that 20% of the world economy is vulnerable to weather conditions (Barrieu & Scaillet, 2010). Weather is one of the most uncontrollable and influential variables within the agriculture sector, becoming increasingly unpredictable as climate change continues to affect global weather patterns. In some cases, extreme weather can cause up to a 40% deficit in crop yields in Russia, potentially devastating a farmer's economic income (Pavlova, Varcheva, Bokusheva, & Calanca, 2014). However, by utilizing various types of insurance, those in the agricultural sector can mitigate their exposure to this financial risk.

Russia's ambitions to become agriculturally self-sufficient and its ban on imported crops have caused its agricultural sector to grow substantially in recent years (Liefert, Serova, & Liefert, 2015). To foster this growth and develop this sector, farmers need insurance policies to protect themselves from risks that are beyond their control, such as weather. Weather derivatives, a type of financial option, can be used to protect farmers from daily fluctuations in temperature and precipitation that catastrophic insurance plans do not shield them from (Chung, 2011). These events have a modest effect over a single day but cumulatively they can have severe effects on a farmer's yield by the end of the growing season. Though weather derivatives have been used to hedge against risks in other countries, Russia has yet to explore this tool and popularize it among its farmers (Esper Group, 2010).

The goal of this project is to create a proof-of-concept weather derivatives pricing system. This system will explore the feasibility of farmers' insurance within Russia using such financial instruments. Farmers will be able to hedge against weather-related risks by trading weather derivative options and to remain financially stable in times of fluctuating weather conditions. To accomplish this goal, we had to meet the following objectives:

Determine the relationship between temperature and crop yields within the Moscow, Krasnodar, and Omsk regions (see Figure 1)

Price weather derivative options

Create an Excel tool to simulate the financial impact of weather derivatives for users

Using Weather Derivatives to Insure Russian Agriculture

To implement a weather derivatives system within Russia, one must understand the rela-

Figure 1. Regions of focus: Left-Krasnodar, Center-Moscow, Right-Omsk.

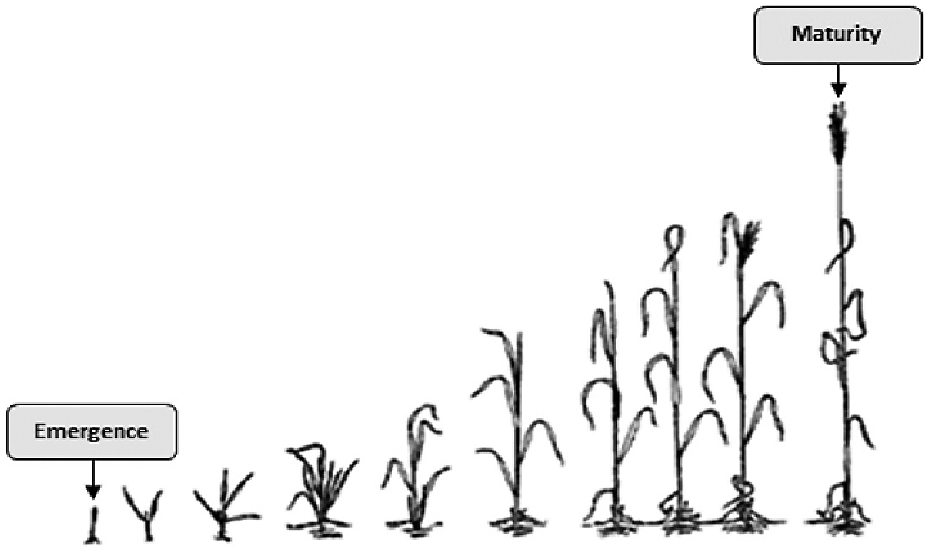
tionship between weather and agriculture and the current measures in place to protect farmers against weather risks. In this chapter, we will explain the concept of a weather derivative to hedge against these risks. Then we will discuss Russia's current agricultural economy and strategies to protect those working in agriculture from losses due to weather events.

Weather risks and mitigation strategies. The associated economic risks tied to weather can be divided into two major groups: high frequency-low risk events and low frequency-high risk events. Low frequency-high risk events, such as tornadoes and hurricanes, have an extreme, immediate impact, costing millions of dollars in damages. High frequency-low risk events are everyday weather phenomena, such as rain and temperature change. These events cause little impact over a single day but cumulatively can cause substantial negative effects. The agricultural sector is especially sensitive to this type of risk, causing weather to have a considerable effect on the economy (Barrieu & Scaillet, 2010).

Governments across the globe have set up various forms of insurance, such as government subsidies or weather derivatives, to protect those in the agricultural sector. The use of subsidies in times of poor harvest, however, is not always ideal or even feasible for less developed countries that cannot generate enough revenue from taxation. Additionally, subsidy compensation is based on a farmer's exact loss, requiring insurers to determine farmer's yields to calculate what compensation is due. This increases costs to the insurer and raises the cost of premiums for those who are insured (Chung, 2011).

Weather derivatives. Weather derivatives offer advantages to both small-scale farmers and corporate agricultural businesses. These derivatives are a type of option with an index-based payout, modeled after predicted future weather conditions over a certain period. The major difference between a weather derivative and subsidy is that the payout for a derivative is based on the specific weather conditions that cause farming loss, while a payout for a subsidy is based on the actual loss itself. Thus, weather derivatives can cover the high frequency-low risk events described above without the need for insurers to determine farmer's exact yields, keeping premium costs lower. However, a substantial amount of meteorological data is required to price the derivative (Chung, 2011).

Around 75% of all weather derivative transactions are based upon temperature predictions while 10% of transactions are based on rainfall (Barrieu & Scaillet, 2010). Temperature-indexed weather derivatives revolve around the concept of Growing Degree Day (GDD), which measures



Days after emergence	12	15	23	28	34	38	41	45	85
Cumulative GDD	170	230	430	500	690	810	895	990	2260

Figure 2. Adapted from 'Growth and Development Guide for Spring Wheat' (Simmons, Oelke and Anderson, 1985).

heat accumulation to predict favorable plant development rates and stages of growth (see Figure 2, Appendix A). The metric below computes the difference between realized temperatures to a baseline temperature, which varies depending on the crop species.

Non-Russian weather derivatives systems. While weather derivatives are still a fledgling concept, being first traded on the Chicago Mercantile Exchange (CME) in 1999, their use is slowly becoming more commonplace within global markets (Barrieu & Scaillet, 2010). The Canadian agricultural insurance market recently introduced weather derivatives to insure against abnormal season temperatures or precipitation levels. After interviewing 397 farmers from Saskatchewan over a period of three years, investigators showed that 307 of these farmers used only traditional agricultural insurance, 37 only used weather derivatives, and 37 used both types of insurance. The study concluded that this wide disparity in weather derivative use is mainly attributed to farmers' lack of "awareness and understanding" of the tool (Van Camp, 2015, para. 5). About half of the participants who did not invest in weather derivatives were not aware that such a tool was available to them. About one-third of these farmers felt they did not have enough knowledge and skill to utilize the derivative (Van Camp, 2015).

In 2003, a Mumbai insurance company implemented weather derivatives for small groundnut and castor farmers in four villages within the Andhra-Pradesh state. The program encouraged farmers to attend educational workshops about the product to inform farmers of this insurance and its benefits, increasing the derivative's approachability. In 2005 after more improvements to the program, "more than 250,000 [sic!] farmers bought weather insurance" (Barrieu & Scaillet, 2010, 7). This pilot project was deemed a major success and inspired many more weather-based insurance schemes across India such as the Weather-based Crop Insurance Scheme (WBCIS) (Ministry of Finance of India, 2017).

One of the main distinctions between the Indian and the Canadian weather derivatives program is the presence of an educational program for the users. Equipped with the knowledge of how these weather derivatives could financially support them, farmers in India widely supported the weather derivatives system. However, those in Canada struggled to see the potential benefits of these tools or were completely unaware of them. Thus, to build a successful and accessible weather derivative system, it is vital to educate the users.

Agriculture in the Moscow, Krasnodar, and Omsk regions. The Russian agriculture sector employs 7.7 million people, or 12% of the total

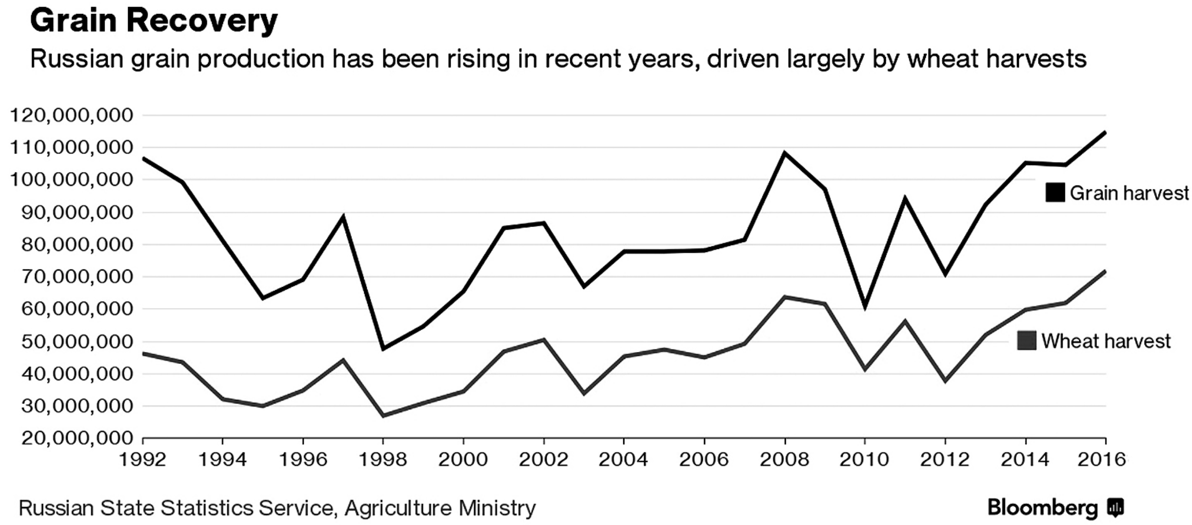


Figure 3. Russian grain production 1992–2016 (Medetsky, 2016).

workforce (British Potato Council, 2006). Most of Russia's land mass is in "risky farming zones," where harvest capacity depends largely on weather conditions. This is exacerbated by global climate change, making weather conditions increasingly unpredictable. Because of the country's geographic span, the overall climate of Russia varies significantly from area to area, allowing different crops to thrive in different temperatures (Country Studies, 1996). These differences not only affect the rate of crop growth, but also the crops' planting and harvest dates, creating a unique set of growing conditions in each region.

Wheat, corn, and potatoes are three of the most widely-grown crops within Russia (Basic Element, 2013). Grains occupy more than 50% of the available cropland, primarily in the form of wheat (Country Studies, 1996). Overall land productivity has recently increased due to a decrease in the price of the ruble and recent favorable growing conditions (see Figure 3) (Medetsky, 2016). These large yields have brought in a substantial income for farmers, but only because of favorable weather conditions. Thus, the agricultural sector is currently a lucrative investment area.

The Moscow, Krasnodar, and Omsk regions provide a representative range of Russian climatic and agricultural conditions. The Moscow region is in the western part of the country. Because of its large population, its local agriculture has a high profile. Krasnodar is the economic center of southern Russia, and 42.8% of its main industries is agriculture-based (Oleynik, 2013). Because of Krasnodar's geolocation by the Black Sea, the region has a longer growing season and ideal weather conditions for plant growth (State's executives of the Krasnodar Region, n.d.). Conversely, the growing conditions in Omsk are not as favorable. Situated on the West Siberian Plain, the annual average temperature in Omsk is around 1.4°C (Climatemp, n.d.). Wheat, corn, and potatoes are grown in all three areas, but each is subject to the region's unique weather conditions.

The shortcomings of the Russian Government subsidies system. Government subsidies are currently used to help farmers in Russia hedge against weather risks (Buckley, 2017). State-issued subsidies have created significant growth within the agricultural sector, but not without complications. Some farmers cannot afford premiums, meet land acreage requirements, or obtain the necessary accounting paperwork to qualify for payments. In the 2012 drought, state compensation was only given to farmers "located in emergency districts… in a manner that was not at all transparent [to the farmers]," while those located in "non-emergency" zones suffered terrible losses as well (Ukhova, 2013, 12). Those who received payment received insufficient amounts in comparison to their actual loss. This underperformance by subsidies has resulted in a general lack of faith in the system (Ukhova, 2013). To work to remedy this, farmers must be able to easily access their method of compensation and understand why they are receiving it. Even with these improvements, subsidies only protect against high-impact events such as droughts. There is still a clear lack of protection against small but continual risk such as temperature fluctuations (Esper Group, 2010).

Conclusions. Weather derivatives can be used to insure farmers against daily fluctuations in temperature, which can have a substantial impact on their yields and wallets. Most of the farmland within Russia is highly sensitive to weather conditions. Though government subsidies have been used in the past to assist farmers in protecting themselves against weather risks, farmers no longer trust this system. Weather derivatives, however, use objective weather data to help farmers compensate for their losses incurred by unfavorable weather conditions. As shown in the Indian and Canadian cases, for the concept of weather derivative to work it must be familiar to farmers, and it is vital that they are educated about this tool's use and benefits. This builds trust and extends the use of an effective weather derivatives system.

Methodology: Developing a Weather Derivatives System

The goal of this project is to create a proof-of-concept weather derivatives pricing system. This system will explore the feasibility of farmers' insurance within Russia using such financial instruments. We created the following objectives to successfully reach this goal:

1. Determine the relationship between temperature and crop yield within the Moscow, Krasnodar, and Omsk regions
2. Price weather derivative options
3. Create an Excel tool to simulate the financial impact of weather derivatives for users.

Determining relationship between temperature and crop yield. Because the pricing of weather derivatives depends upon GDDs that are crop-specific, we selected 3 regions and 3 specific crop types for the construction of derivatives. We identified corn, potatoes, and wheat (spring and winter) as some of the most common crops in Russia and the Moscow, Krasnodar, and Omsk regions as areas representing a spread of weather conditions. We gathered each crop's baseline temperature for its GDD calculation, its planting dates, and its harvest dates. Using these dates and temperatures, we accurately gauged the temperatures these crops experience within a growing season. We calculated the mean cumulative GDD experienced by each crop within the Moscow, Krasnodar, and Omsk regions from the years 1996 to 2015 with data from the meteo.ru (RIHMI-WDC) weather database and collected regional crop yield statistics from Knoema, another online database (see References). Using Microsoft Excel, we developed a database of these temperatures and implemented an ordinary least squares regression technique to quantify the relationship between cumulative GDD over the growing period and crop yield.

Pricing weather derivatives. To price the derivatives, we surveyed various pricing methods. After reviewing literature by Sun and van Kooten (2015); Groll, López-Cabrera, & Meyer-Brandis (2016); Taylor & Buizza (2006); Chung (2011); Alaton, Djehiche, & Stillberger (2002); Barrieu & Scaillet (2010); and Consedine (2000), we chose the historical burn analysis method, which takes the average historical GDD as the expected GDD for future years (see Appendix A). This technique was chosen because of its ability to accurately model these future GDD values, the accessibility of the data needed for this method, and the ability to conduct the necessary mathematical processes in a familiar format such as an Excel spreadsheet.

Creating simulation tool. To visually represent the results of this project and demonstrate the potential impact of this weather derivatives system, we created a weather derivative simulation tool in Visual Basic for Excel. This tool calculates the potential losses a farmer faces by interfacing with the GDD/yield relationship model. The farmer inputs his/her farm size, crop type, and location. His/her projected yield for the upcoming year is then calculated by utilizing the appropriate GDD/yield model, the projected GDD based on his/her region, and the size of his/her farm. This yield is then multiplied by the estimated worth of his/her crop, data gathered from Bloomberg, converting his/her potential profit to a monetary value.

Based on the GDD/yield model, the tool also estimates potential economic loss if the weather varies from the expected GDD. A derivative is then constructed using the chosen tick size. The derivative's payoff can be compared to a farmer's potential loss, showing its potential effectiveness as a form of insurance. The tool draws upon values from the database mentioned in Section 3.1. Because all the data inputs (excluding those provided by the user) are contained within Excel spreadsheets, the tool can be easily updated to include more recent information or different areas and crops, expanding it to become a more encompassing and accurate tool.

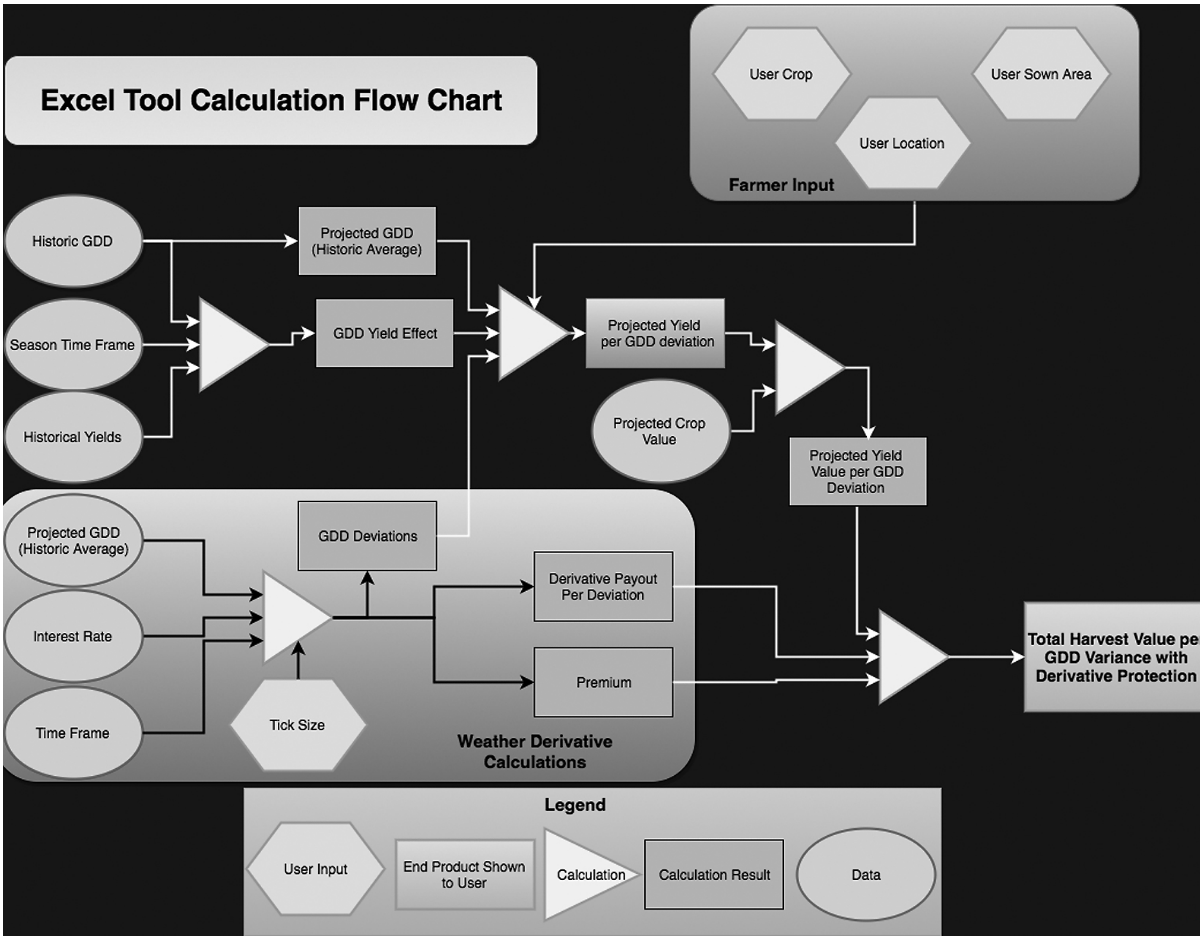


Figure 4. The main components of weather derivative simulation tool.

Conclusions. The cumulative application of our methods is showcased in the simulation tool. The tool can evaluate the GDD/yield relationship for each region and crop, the predicted GDD values for future years, and the potential profit or loss with weather derivative use for a specific user. This allows the user to visualize the effect of a weather derivative and its potential as an insurance measure. Additionally, the tool is easily modifiable, allowing it to remain relevant and open for modification while further developments take place in this research field. By using this tool, those who are interested in developing derivative-based insurance can also test their own research methods and display these techniques to their target users.

Results and Discussion

After initial poor results in our regression analysis for cumulative GDD and crop yield, we found there were large flaws in the methods in which we were processing and interpreting our collected data. We then developed a strategy to correct these flaws to pre-process our data to eliminate trends that were contaminating our results. This leads to more accurate results, producing a clearer relationship between the two variables. When pricing the derivative, the historical burn analysis generated high-quality GDD predictions and generally low premiums for the farmers. Both the regression and the pricing calculations were implemented in our Excel simulation tool that is both flexible for those who wish to build upon it and approachable for farmers who wish to use it.

Determining Relationship between Temperature and Crop Yield

The regression between temperature and crop yield initially yielded weak results and no clear or logical relationship has been obtained at that point (see Figure 5). After discussing the quality of our data, we isolated the causes of this weak regression result to two factors:

1. Qualitative growing season data
2. Skewed yield data

When collecting harvest and planting dates, we found that the data was extremely qualitative, described as "early May", "mid-September", etc. This is reasonable for a farmer who plants when the soil

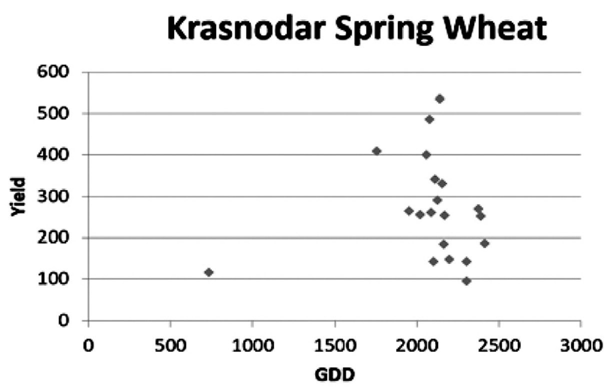


Figure 5. Krasnodar spring wheat GDD/yield before data pre-processing (1996–2015)

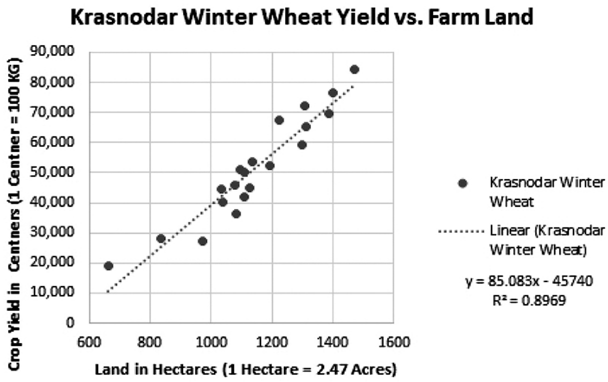


Figure 6. The trend between acreage and crop yield for Krasnodar winter wheat (1996–2015).

Table 1

Yield Per Area vs Cumulative GDD Regression R 2 Values

Crop	Moscow	Krasnodar	Omsk
Corn	0.276	0.201	0.137
Potato	0.026	0.247	0.195
Spring Wheat	0.211	0.428	0.26
Winter Wheat	0.083	0.259	0.057

Figure 7. Krasnodar spring wheat and Moscow corn regression results (1996–2015).

is deemed ready, but not sufficient for quantitative analysis. To accurately model these decisions, we further researched the favorable planting conditions for our four crops. Then, based on this information, we created an algorithm to search through the temperature database and select a planting day that meets these conditions.

Each crop has its respective GDD criteria to meet to reach its planting date. However, GDD is not the only factor used. The typical growing season for our chosen crops covers a period of three months. Thus, our algorithm selects a planting date that satisfies the GDD requirements within this time range. If this criterion was not met during this period, the end of the time interval was selected as the planting date. This method of selecting planting dates creates a more accurate picture of actual GDD, giving us stronger models to predict crop growth. Harvest dates, on the other hand, remain relatively stable from year to year and do not require such attention.

We then realized that our collected yield data had varying amounts of total acreage per year contributing to this yield. An increased total acreage was resulting in an increased total yield for that year, i.e. causing a linear trend within the data (see Figure 6). Thus, to isolate the effects of GDDs on crop yield, we converted the raw yield data into yield per recorded acreage. The regression analysis then produced varying results (see Figure 7 and Table 1). Thus, this relationship can be used to approximate how a predicted change in cumulative GDD in each region will affect the yield results for each crop, demonstrating the farmers' potential loss in yield. This is the first step in showcasing

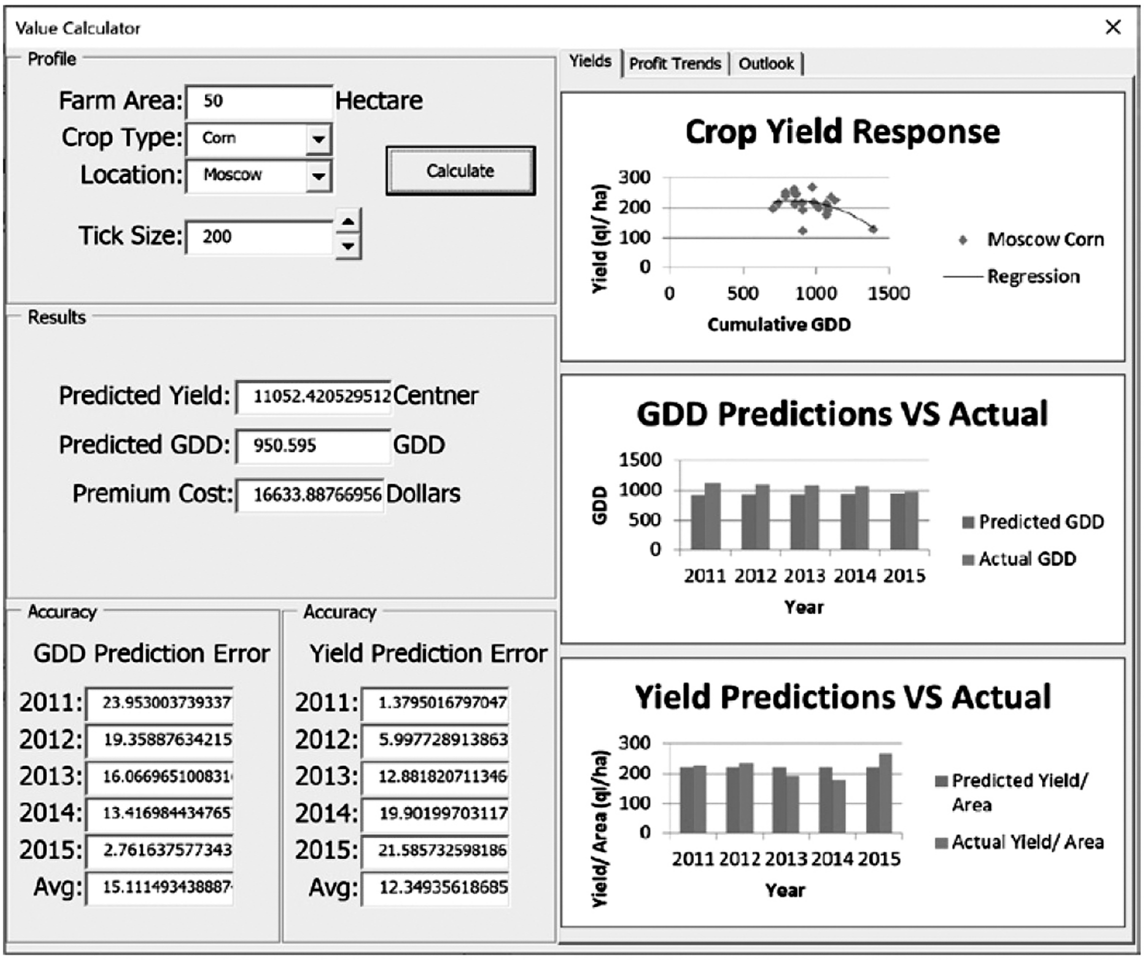


Figure 8. Weather derivative simulation tool interface.

to them how the purchase of weather derivatives can compensate for this projected loss.

Pricing Weather Derivatives

Following the formulas for pricing the weather derivatives, the farmer profits whenever the GDD hits one of two appropriate points (see Appendix A). However, it was not clear how to adjust these pricing parameters so that farmers with a larger amount of farmland and a greater economic loss from poor weather conditions would be able to buy a weather derivative to collect a larger payout. In other words, we could not establish a relationship between farm size and premium. Therefore, we added tick size as a user input for our simulation tool.

Creating Simulation Tool

The final deliverable of our project is an easy-to-use tool that compiles all our work and demonstrates the effectiveness of weather derivatives to farmers, while also serving as a stepping stone for a practical implementation of this project. The tool performs situation-specific calculations based upon profile information provided by the user, e.g. crop type, location, farm size, and tick size (see Figure 8). Using this information as a basis for our parameters, the tool draws from an Excel database to calculate the GDD/yield relationships, predicted GDDs and yields, the potential profit/loss of the farmer, and the price of the weather derivative. The farmer is then able to see his/her potential loss under various circumstances.

The program offers a large amount of flexibility in terms of upkeep, update potential, and data management. Data can easily be added to the Excel database for further processing as more weather and yield data is collected. The tool itself can easily be used by those with basic familiarity with Microsoft Office Products. The program's functionality demonstrates the potential effectiveness of utilizing weather derivatives for farming insurance and serves as a flexible and scalable tool that can generate further interest in the development of a weather derivatives program.

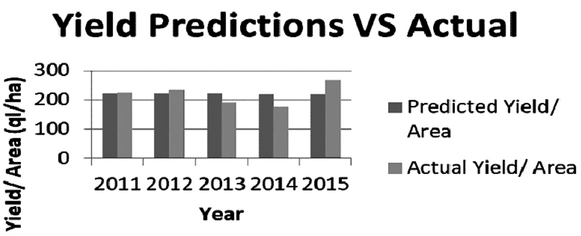


Figure 9. The accuracy of Predicted Yield Values.

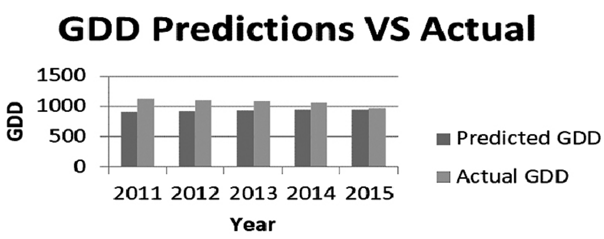


Figure 10. The accuracy of Predicted GDD Values.

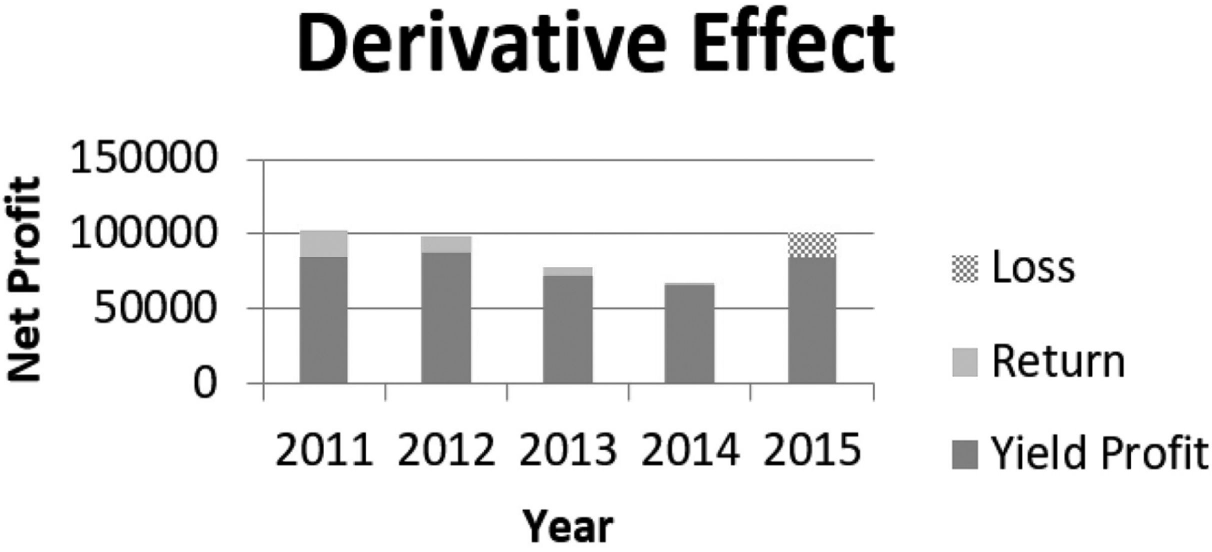


Figure 11. Profits with yield and weather derivative use.

Testing Simulation Tool

To determine the accuracy of the simulation tool and whether weather derivatives are an effective hedging tool for farmers, we added testing code to the tool. This test code takes the last 5 years of the database (years 2011–2015) and treats them as future years. For each of these years, we predict the GDD, the crop yield, and produce a derivative option. The actual yield and GDD are then compared to what was predicted to determine accuracy, and the farmer's profit is compared to the derivative return to determine its effectiveness. As each past year is tested, its information is added back into the database for the next test calculation. The results show that as more data is added to the model, it becomes more accurate (see Figure 9 & 10). It also shows that with adjustment of the tick size, weather derivatives can help considerably to cover farmer's losses (see Figure 11).

Advantages of the Simulation Tool

This simulation tool allows individuals to quickly visualize the potential benefits of utilizing weather derivatives as insurance. A farmer, or someone acting as a farmer for academic research, can input information that reflects their current economic position, and then gauge how effectively weather derivatives can mitigate their economic risks. In terms of development, it allows researchers to determine the efficiency of weather derivatives and adjust parameters as necessary when working towards a market implementation. For users, the tool's ability to easily convey the savings delivered by a derivative should generate popular interest in the product. The creation of this tool will hopefully spur the development of derivative-based insurance systems throughout Russia to further boost the agricultural sector development.

Conclusions and Recommendations

From our project work, we have compiled a list of recommendations for the further development of this weather derivative tool. Ultimately, we recommend:

Testing the tool with real users

Promoting the tool amongst real users

Conducting laboratory experiments to determine the effect of precipitation on yield and create precipitation-based derivatives

Optimizing pricing parameters

Evaluating and applying other pricing techniques

Trading the weather derivatives on a local exchange trading system (LETS)

Testing with Real Users

To confirm the effectiveness and reliability of this tool, it is imperative that actual farmers test it. These farmers would complete surveys and/or take part in focus groups to evaluate the ease of use of this tool. Additionally, these farmers can judge the robustness of the constructed models. Users would record their actual crop loss versus their predicted loss and their actual compensation from the derivatives. The differences in the actual conditions and projected conditions would then be used to create more accurate models, creating a more beneficial tool in the future.

Promoting the Tool

Once the tool has been sufficiently tested, it is important that farmers are aware that this tool exists. As described above, many Canadian farmers did not know that weather derivatives existed or how they could be utilized (Van Camp, 2015). Thus, we recommend that our tool is promoted in a marketing campaign. This promotion would involve researching the methods of communication that are most valuable to farmers (e.g. publications in an agricultural magazine, workshops like those in the India system, word-of-mouth, etc.) and then promoting through these methods. The farmers will never be aware of how this tool can help them if they are never aware of the tool itself.

Testing the Effect of Precipitation and Constructing Precipitation-based Derivatives

One issue we encountered during the development of the GDD/yield model is that, even with the preprocessing of data, cumulative GDD is not the only factor that determines crop growth. As evident in Figure 8, some years it experiences a similar GDD but vastly different yields. Precipitation also plays a key role in crop development. With global changes in precipitation patterns, it is important to factor in more than just temperature into our yield response model. Thus, we recommend conducting a laboratory experiment that analyzes the effect precipitation has on overall yield for these crops. This experiment would expose crops to the same cumulative GDD, but change the amount of water each plant receives and document each crop's growth rates. A similar experiment should also be conducted that maintains constant water levels and varies the GDDs.

Comparing the results of each of these tests would reveal which factor is more critical for the growth of different crops. A similar precipitation/yield model could be constructed so that farmers can visualize how future rainfall predictions will affect their crops. Weather derivatives based on a cumulative rainfall index can also be priced. This will allow farmers to pick between a GDD or a precipitation derivative, depending on whichever is more unpredictable and/or influential in their region.

Optimizing Pricing Parameters

Further research must be done to optimize the pricing parameters to ensure that farmers' premiums are affordable to the farmer and that these payouts provide substantial compensations. For example, the strike values of the weather derivative are currently set at 0.2 standard deviations away from the mean cumulative GDD values. By setting the strike values at a larger standard deviation away, we decrease the cost of the premium, but also decrease the likelihood of receiving a payout from the derivative. Thus, a balance must be found between the initial premium cost and meaningful levels of compensation.

Evaluating Other Pricing Methods

Finding the best methods to price weather derivatives is an open research problem. As stated before, we selected the historical burn analysis because the data needed for processing was accessible and the technique proved to be effective in previous research papers. The mathematical concepts presented were also easy to grasp and implement by our team in Excel within a limited timeframe.

Currently, more accurate methods of pricing exist, even if they were not feasible for our team to calculate. For example, Taylor and Buizza (2006) use ensemble forecasting to create their weather prediction model with data provided by the European Centre for Medium-range Weather Forecasts (ECMWF), a source which we did not have access to. With higher-fidelity forecasting models, more accurate derivative pricing will ensue and more protection will be provided to the farmers. Because weather prediction continues to be uncertain, we recommend a more comprehensive comparison of weather derivative pricing that encompasses techniques outside of those presented here. This will either affirm the accuracy of our methods or provide even more accurate pricing methods.

Trading Weather Derivatives

Most weather derivatives are currently traded on the market using over-the-counter (OTC) transactions, meaning they are not traded on formal exchange systems like NASDAQ or Dow Jones but privately negotiated between two parties (Investopedia, n.d.). We did not pursue research into bringing the derivatives to a real-world market due to lack of time for the project. Eventually, this weather derivatives system should be brought out of academia and into the real world. We recommend further research into trading derivatives on an online local exchange trading system (LETS) so that contracts can be easily bought and sold all around the world. Additionally, all derivative transactions could take place utilizing Blockchain technology (Figure 12), eliminating the need for clearing houses as well as third-party security issues. This would also decrease costs to users and increase their profits (Iansiti & Lahkani, 2017).

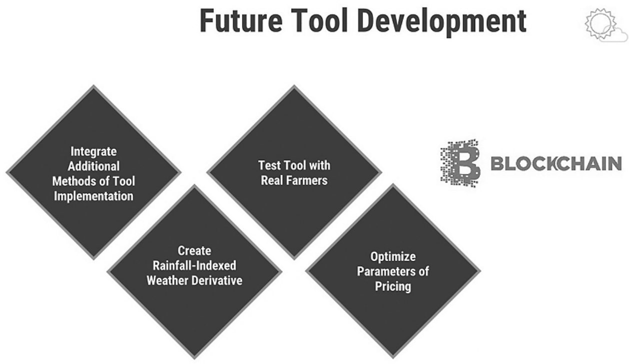


Figure 12. Further developments of the author's weather derivatives system.

Conclusion

With global climate change altering weather patterns, Russian farmers need protection from everyday weather events that will negatively affect their crop yields. This type of protection is not currently offered through traditional methods of agricultural insurance or government subsidies and furthermore, Russian farmers have a lack of faith in these products. Using weather derivatives, these farmers should be able to hedge these risks at an affordable premium price. To build a weather derivative simulation tool, our team constructed a model that demonstrates the relationship between cumulative GDD within a growing season and crop yield for corn, potatoes, and wheat in the Moscow, Krasnodar, and Omsk regions. We were then able to price weather derivatives, displaying these results and models on the Excel simulation tool. This tool can demonstrate how predicted GDDs will affect farmers' yields and how they can protect themselves from potential economic loss and thus boost popular interest in a weather derivatives system in Russia.

References

Alaton, P., Djehiche B., & Stillberger, D. (2002). On modeling and pricing weather derivatives. *Applied Mathematical Finance, 9*(1). doi:10.1080/13504860210132897.

Basic Element. (2013, December 9). The future of farming in Russia. *Farmers Weekly*. Retrieved from www.basel.ru/en/articles/farming_09_12_13.

Barrieu, P., & Scaillet, O. (2010). A primer on weather derivatives. In *Uncertainty and Environmental Decision Making: A Handbook of Research and Best Practice* (pp. 155–175). Retrieved from http://1cellonelight.com/pdf/weatherCambridge.pdf.

British Potato Council. (2006, August). *Target market report for the export of GB seed potatoes*. BPC Target Market Report Russia. Retrieved from https://potatoes.ahdb.org.uk/sites/default/files/publication_upload/BPC%20OMIS%20Report%20Russia.pdf.

Buckley, N. (2017, April). Russian agriculture flourishes amid sanctions. *Financial Times*. Retrieved from https://www.ft.com/content/422a8252–2443–11e7–8691-d5f7e0cd0a16.

Climatemp (n.d.). Omsk Weather and Temperature. Retrieved from http://www.omsk.climatemps.com.

Chung, W. (2011). Evaluating weather derivatives and crop insurance for farm production risk management in southern Minnesota. Master's thesis, University of Minnesota. Retrieved from https://conservancy.umn.edu/bitstream/handle/11299/119325/Chung_umn_0130E_12390.pdf?sequence=1&isAllowed=y.

Consedine, G. (2000). *Introduction to Weather Derivatives*. CME Group. Retrieved from https://www.cmegroup.com/trading/weather/files/WEA_intro_to_weather_der.pdf.

Country Studies. (1996). Agriculture. Country Studies. Retrieved from http://countrystudies.us/russia/60.htm.

Esper Group. (2010). NO CLIMAT but weather. *Esper Times*. Retrieved from http://www.esper-group.com/files/4/Esper-Times-_30-eng.pdf.

Glenn, J. (2017). Euro vs. US Dollar. Investor Guide. Retrieved from http://www.investorguide.com/article/15564/euro-vs-us-dollar-d1412.

Global Investment & Business Center. (2013). *Russia Regional Economic and Business Atlas* (Vol. 2). Washington, DC: International Business Publications. Retrieved from https://books.google.ru/books?id=ZEe8BQAAQBAJ&printsec=frontcover&hl=zh-CN&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false.

Groll, A., López-Cabrera, B., & Meyer-Brandis, T. (2016). A consistent two-factor model for pricing temperature derivatives. *Energy Economics, 55*, 112–126. Retrieved from http://www.sciencedirect.com/science/article/pii/S 0140988316000104.

Iansiti, M., & Lakhani, K. (2017, January–February). The truth about blockchain. *Harvard Business Review*. Retrieved from https://hbr.org/2017/01/the-truth-about-blockchain.

Investopedia. (n.d.). Over-The-Counter-OTC. Retrieved from http://www.investopedia.com/terms/o/otc.asp.

Knoema. (2015). Production and Sale of Agricultural Products, 1990–2015. Russian Agriculture Yields. Retrieved from https://knoema.com/atlas/Russian-Federation/Krasnodar-Krai/topics/Agriculture/Production-and-sales-of-agricultural-products-Gross-harvest-of-crops/Yield-of-winter-wheat.

Liefert, W. M., Serova, E., & Liefert, O. (2015). Russia's economic crisis and its agricultural and food economy. *Choices*, *30*(1), pp. 1–6. Retrieved from http://www.choicesmagazine.org/magazine/article.php?article=78.

Medetsky, A. (2016, October 6). Russia Becomes a Grain Superpower as Wheat Exports Explode. Retrieved from https://www.bloomberg.com/news/articles/2016–10–06/russia-upends-world-wheat-market-with-record-harvest-exports.

Ministry of Finance of India. (2017). *Government Sponsored Socially Oriented Insurance Schemes*. Department of Financial Services, Government of India. Retrieved from http://financialservices.gov.in/insurance/gssois/wbcis.asp?pageid=2.

Pavlova, V. N., Varcheva, S. E., Bokusheva, R., & Calanca, P. (2014). Modelling the effects of climate variability on spring wheat productivity in the steppe zone of Russia and Kazakhstan. *Ecological Modelling*, *277*, 57–67. Retrieved from http://www.sciencedirect.com/science/article/pii/S 030438001400057X.

RIHMI-WDC, Russia, CDIAC, USA. Daily Data. Air Temperature and Precipitation. (n.d.). Retrieved from http://meteo.ru/data/162-temperature-precipitation#доступ-к-данным.

Simmons, S. R., Oelke, E. A., & Anderson, P. M. (1985). *Growth and Development Guide for Spring Wheat*. St. Paul, MN: University of Minnesota Agricultural Extension Service. Retrieved from https://conservancy.umn.edu/handle/11299/165834.

State's executives of the Krasnodar Region. (n.d.). *Agricultural of Krasnodar Region*. Retrieved from https://www.krasnodar.ru/en/content/453/show/32845.

Sun, B., & van Kooten, G. (2015). Financial weather derivatives for corn production in Northern China: A comparison of pricing methods. *Journal of Empirical Finance*, *32*, 201–209. Retrieved from http://www.sciencedirect.com/science/article/pii/S 0927539815000328.

Taylor, J., & Buizza, R. (2006). Density forecasting for weather derivative pricing. *International Journal of Forecasting*, *22*, 29–42. Retrieved from http://www.sciencedirect.com/science/article/pii/S 0169207005000634.

Ukhova, D. (2013, September). After the Drought: The 2012 Drought, Russian farmers, and the challenges to adapting to extreme weather events. Oxfam Case Study. Retrieved from https://www.oxfam.org/sites/www.oxfam.org/files/cs-russia-drought-adaptation-270913-en.pdf.

van Camp, M. (2015). Weather derivatives to mitigate weather risk to crops. Country Guide [Paragraph 3, 4 and 5]. Retrieved from https://www.countryguide.ca/2015/11/23/weather-derivatives-to-mitigate-weather-risk-to-crops/47647.

Acknowledgements

Our team would like to thank the following individuals for their contributions to the completion of our project:

Worcester Polytechnic Institute and The Financial University under the Government of the Russian Federation for providing us with the means and opportunity to complete this project.

Alexander Ilyinsky, Dean of the International Finance Department at the Financial University, for his guidance and for planning the foundations of the project as our sponsor.

Tatiana Goroshnikova, Deputy Dean of the International Finance Department at the Financial University, for her weekly advice and guidance.

Svetlana Chugunova for organizing our stay and accommodations while in Russia.

Oleg Pavlov, Stephan Sturm, Gbetonmasse Somasse, Michael Radzicki, Michael Johnson, and Anton Losev, for sharing the knowledge of their respective fields during our interviews and team meetings.

Svetlana Nikitina for coordinating with The Financial University under the Government of the Russian Federation to make this project possible and for her continual guidance throughout the project as our advisor.

About the authors

Eric Carkin, student of Robotic and Electrical engineering at WPI, ecarkin@gmail.com

Stanislav Chekirov, student of International Finance Faculty at Financial University, chekirovstas@gmail.com or 142195@edu.fa.ru

Anastasia Echimova, student of International Finance Faculty at Financial University, echimova.anastasia@gmail.com or 164615@fa.ru

Caroline Johnston, student of Mathematical Sciences major and Computer Science minor at WPI, Ceej39@gmail.com or cmjohnston@wpi.edu

Congshan 'Shannon' Li, student of Architecture Engineering and Mechanical Engineering minor at WPI, Shan1143148746@gmail.com or Cli8@wpi.edu

Vladislav Secrieru, student of International Finance Faculty at Financial University, vlasec08@gmail.com or 144077@edu.fa.ru

Alyona Strelnikova, student of International Finance Faculty at the Financial University, lkastrelnikova@gmail.com or gimnaz23@kubannet.ru

Marshall Trier, student of Robotic engineering at WPI, Marshalltrier@gmail.com or mtrier@wpi.edu

Vladislav Trubnikov, student of International Finance Faculty at Financial University, WolveVlad@gmail.com or unclelad777@mail.ru

Appendix

Appendix A. Relevant Equations

A Growing Degree Day (GDD) is defined as

$$GDD_{i,n,c} := \frac{T_{max,i,n} + T_{min,i,n}}{2} - T_c\,,$$

where $T_{max,i,n}$ and $T_{min,i,n}$ are the maximum and minimum recorded temperatures, respectively, for day, i and year, n; and T_c is the base temperature for crop, c.

Cumulative GDD is defined as

$$\sum_{i=s}^{q} GDD_{i,n,c}\,,$$

where s and q are the start and end dates of the growing season, respectively.

The expected payout for a weather derivative with low GDD or high GDD is defined as

$$E_{p,LOW} = D\sigma\left[\phi(n) + n\Phi(n)\right]$$

$$\text{or } E_{p,HIGH} = D\sigma\left[\phi(m) - m + m\Phi(m)\right],$$

where D is the tick size (dollar value per unit of GDD), μ is the mean value of GDD's, σ is the standard deviation of the GDD's, ϕ is the PDF of the standard normal distribution, Φ is the CDF of the standard normal distribution, and

$$n := \frac{K_1 - \mu}{\sigma},\ m := \frac{K_2 - \mu}{\sigma},$$

where K_1 is the strike value for the low GDD value, and K_2 is the strike value for the high GDD value (see Sun and van Kooten (2015) for derivation). The dollar is used as the choice of currency in the tick size because of its historic stability (Glenn, 2017). Thus, the price or payout of an option will not fluctuate due to inflation.

The price (premium) of the option is defined as

$$c = e^{-r(u-v)} E_p,$$

where c is the premium that hedgers pay for the contract, r is a risk-free periodic market interest rate, v is the date that contract was issued/purchased, and u is the date the contract was claimed/expiration date. E_p is the expected payoff based on predicted or historic mean value of temperatures (see Sun and van Kooten for the derivation).

The actual payout is defined as,

$$p(x)_{farmer} = \begin{cases} D(K_1 - x),\ x \le K_1 \\ 0, K_1 < x < K_2 \\ D(x - K_2),\ x \ge K_2 \end{cases},$$

where x is the realized cumulative GDD.

In a historic burn analysis, the expected payout is set equal to the average historical weather conditions. In the case of GDD, it is defined as

$$\mu := \frac{\sum_{j=1}^{n} \sum_{i=s}^{q} GDD_{i,j,c}}{n}.$$

In the derivative tool, the interest rate, contract length, risk loading factor, and m and n values were fixed. Additional information (including a downloadable version of the tool) can be found on our website at https://sites.google.com/view/russiaweatherderivatives/home?authuser=0

Погодные деривативы в России: страхование фермеров от колебаний температуры

Eric Carkin[1], Станислав Чекиров[2], Анастасия Екимова[2], Caroline Johnston[1], Congshan Li[1], Vladislav Secrieru[2], Алена Стрельникова[2], Marshall Trier[1], Владислав Трубников[2]

Этот проект предлагает использовать погодные деривативы — тип финансового инструмента с выплатой на основе погодных условий — как метод для российских фермеров для подстраховки от ежедневных колебаний температуры. Мы создали инструмент моделирования производных погоды в Microsoft Excel, который вычисляет влияние температуры на урожайность, а затем анализирует, как возврат производных погоды может потенциально компенсировать потерю урожая. На основе этого инструмента мы разработали ряд рекомендаций, которые помогут внедрить эту систему защиты реальными пользователями.
Ключевые слова: производные погоды; колебания температуры; хеджирование; потери урожая
JEL classification: G13, G17, Q18

[1] Worcester Polytechnic Institute, MA, USA
[2] Финансовый университет, Москва, Россия

Review of Business and Economics Studies
2018, Vol. 6, No. 1, 43–55

DOI: 10.26794/2308-944X-2018-6-1-43-55

Does Social Inequality Stimulate the Economic Growth? (On the examples of the chosen developing countries)

Alina Pukhaeva[1], Elena Miroshina (Silantieva)[2]
Financial University

[1]International Finance Faculty, Financial University, Moscow; alexandra.puhaeva@gmail.com
[2]Ph.D. in economics, Associate Professor, Department of World Economy and International Finance, Financial University, Moscow; EAsila@yandex.ru

Abstract
The article critically examines the concept of social inequality, and suggests ways to determine it against the background of a wide range of factors that determine inequality among the richest and poorest countries. It also summarizes the inequalities between the three groups of countries by comparing some macroeconomic indicators of socio-economic inequality. We then checked for a linear relationship between the two quantitative variables. Using World Bank data and Reports of the United Nations on human development, we conducted an analysis of individual countries taken from three groups of countries (a total of thirty countries), for the period from 1990 to 2017. After a statistical analysis, we proved that inequality slows down economic growth.
Keywords: Inequality, income, GDP per capita, poverty, HDI
JEL classification: I31, J11, J24, D63

Literature Review

It is a common knowledge that the modern world in which we live is unfair and unequal. In general, inequality is regarded to be socio-economic, which in terms means that it is based on income. However, this is not a single measure of inequality but it is rather closely associated with social inequality. Researchers agreed to define social inequalities as differences in income, resources and status within and between individuals. These inequalities are maintained by those in power mostly via institutions. Differences in income distribution matter for a number of reasons, due to the fact that they truly reflect social injustice as well as represent the levels of happiness, according to Ortiz and Cummins (2011).

Despite the ongoing debates within the literature, there is a common belief that global inequality exists and that there are groups across the world that hold more wealth than others and those groups who live in poverty. There is numerous available and descriptive statistical data that presents vast income inequalities across the world, most of them can be found on some official reports of United Nations.

Inequality among people across the world is clearly presentable and is shown in numerous datasets. Thereafter, there is a common knowledge among scientists that inequality in the worldwide scale is vast. However, according to the Milanovic (2007), the way in which inequality is traveling is under the dispute. Based on the report of United Nations (2008) some researchers state that global inequality has been increasing in the past decades both nationally and internationally, with the increasing number of people living in countries in which income differentials are on rise. Interestingly, these increases are not all associated with high-income countries as the current trend is directed toward rising inequality within countries (Firebaugh, 2004). By contrast, other sources of data based on reports of United Nations (2009) depicts that living standards are increasing in some countries and poverty is decreasing as well. According to Milanovic (2010), measurement problems abound, assumptions may be made when

Table 1
Global comparisons in inequality measures for the year of 2015

Country	Human Development Index (HDI) values	Education levels (measured by school enrollment ratio)	GNI per capita, 2015 PPP (US dollars)
Austria	0.893	100.02	43,609
Norway	0.949	112.99	67,614
United Kingdom	0.91	127.81	37,931
China	0.738	94.29	13,345
Turkey	0.767	102.49	18,705
Mexico	0.762	90.55	16,383
Mali	0.442	41.31	2,218
Madagascar	0.512	38.44	1,320
Uganda	0.493	23.24	1,670

Source: Compiled by the authors on the basis of the data of Human Development Reports (http://hdr.undp.org/en) and World Bank dataset (https://data.worldbank.org).

calculations are being done and the selection of some countries within the measurement of overall datasets and trends also seem to skew and distort the larger picture.

The most important question for us is: how to measure social inequality and on the basis of some absolute values do analysis? By looking at opinions of different authors, I have summarized several methods of calculating it. First, according to Milanovic (2007), socio-economic inequality can be measured by comparing the average incomes of different countries. However, this approach, obviously, fails to measure inequality within countries. Second, global inequality can be measured in terms of individual incomes. Again, this is very meticulous work as it involves carrying out surveys, and, thus, is considered to be problematic and quite inaccurate. Third, the most appropriate and accurate method of measuring social inequality from my point of view, is the Human Development Index (HDI), which takes into account three dimensions: Health, Education and Living standards.

Overview of the Global Comparisons in Inequality

Let's conduct the analysis of global comparisons in inequality measures based on the measurement scales of selected countries. According to figures of the World Population Ageing report by the United Nations (2015), there is a distinct depiction of large disparities and, subsequently,

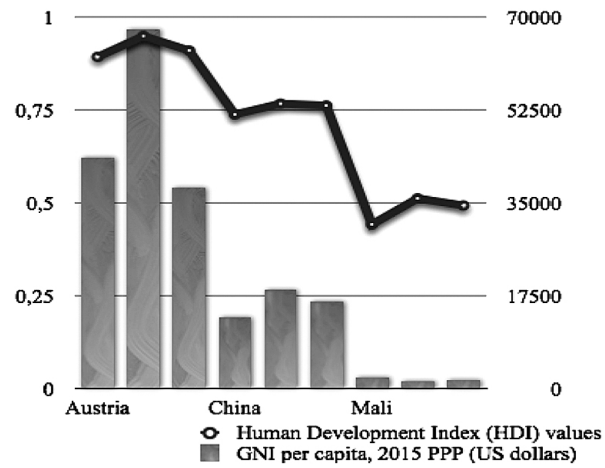


Figure 1. Trends in global comparisons in inequality measures in 2015.

inequality among many countries based on variety of metrics. Table 1 presents some of these inequalities. Due to the availability of the data, we considered only a period of 2015 instead of 2016 or 2017.

Based on the analytical data presented in Table 1, we can state that the figures of Human Development Index (HDI) are monotonically decreasing with respect to the level of development of countries and are positively associated with values of GNI per capita. Thus, for the high-income countries i.e., Austria, Norway and the UK, values of HDI were approximately 0.92 in 2015 since this category of countries are regarded to be highly developed. The average coefficient of HDI would be expected to be equal to unity if the countries

Table 2
Poorest and richest in the world in 2016, measured according to GDP (US dollars)

Poorest country	GDP per capita	Richest country	GDP per capita
Mali	779.9	Luxembourg	100,573.1
Burundi	285.7	Austria	44,676.4
Mozambique	382.1	Qatar	59,324.3
Niger	364.2	Denmark	53,549.7
Uganda	580.4	Norway	70,911.8

Source: Compiled by the authors on the basis of the data of the World Bank dataset (https://data.worldbank.org).

were perfectly developed. Smaller-income countries such as China, Turkey and Mexico, exhibited lower figures of HDI reflecting the fact that they yield smaller average of 0.76 and are regarded as developing countries. As for the countries of the third world (Mali, Madagascar and Uganda), the values of HDI are the lowest and account for only an average of 0.48. This is the result of two factors: countries of the third world are associated with low development capacity of human beings and reflect shortage of capital.

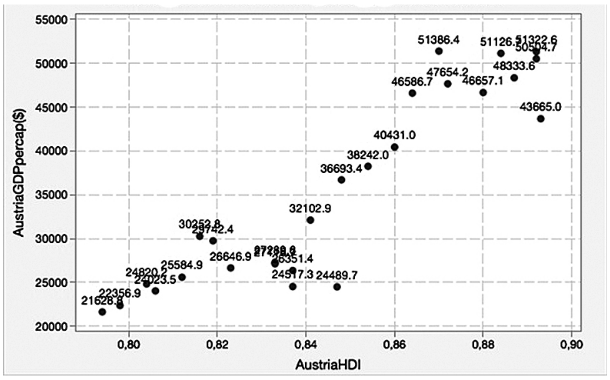
With regard to Education Levels measured by school enrollment ratios, they also show monotonically decreasing pattern. Because of the fact that education levels are measured by school enrollment ratios, the average coefficient of it may not be expected to equal one hundred. In particular, since it may reflect late enrollment as well as early enrollment and repetition, the total enrollment could be expected to exceed the population of the age group which correspond to the official level of education—resulting in ratios greater than 100 per cent. Overall, the relation between high-income countries and school enrollment ratios is clear-cut and allows to state that high coefficients of enrollment ratios are associated with high levels of HDI and GNI per capita. The average ratio of school enrollment for high income countries i.e., Austria, Norway and UK, accounts for 113.6 and a little bit less for countries such as China, Turkey and Mexico yielding an approximate figure of 95.78. By contrast, the poorest countries (Mali, Madagascar and Uganda) demonstrate that enrollment ratios are much lower compared to those of developed countries and are equal to an average of 34.33.

Measuring Global Income Inequality

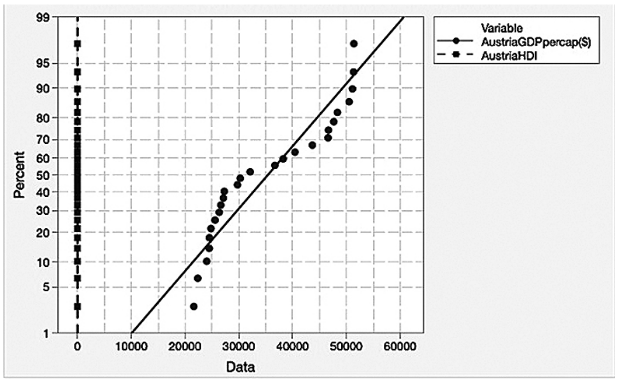
Before proceeding to the examining global income inequality, it is important to consider the point of view of Sen (1999) regarding measuring inequality. He outlines that standards of living are important and play a crucial role in measuring this figure. This approach is connected to the poverty as well as justice and freedom. However, in his empirical work there are measurement errors, which he pointed out, related to the correct evaluation of the standards of living. Thus, in order to solve them out, indicators of social inequality such as gender, hunger and development were taken into consideration and included in the regression analysis of his empirical work. Another indicator of the degree of development is said to be well-being as high-income countries do not always signify the highest level of inequality.

According to Ortiz and Cummins (2011), it is possible to examine global income inequality based on Gross Domestic Product in absolute values, comparing the poorest and richest countries in order to show the severity of current global inequality based on the selection of several countries and presented in Table 2.

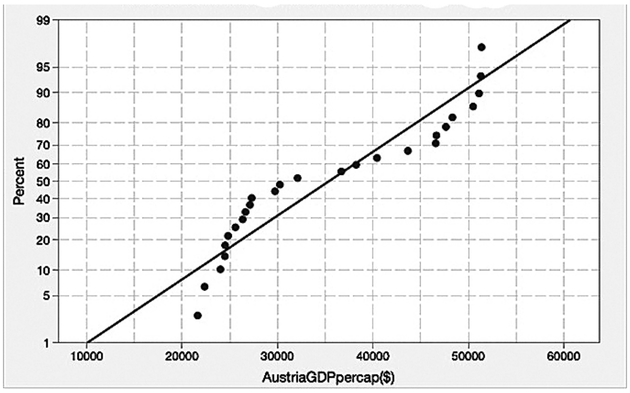
Based on the data extracted from World Bank dataset and presented in Table 2, it is straightforward to suggest a possible view point that GDP per capita measured in US dollars is a further demonstration of the level of inequality experienced today as the table demonstrates. Those countries with the highest levels of poverty are located on the left hand-side, according to 2015 measures, and the richest countries on the right hand-side. The table illustrates some interesting features by comparing the amounts of GDP per capita in two different groups of countries with regard to the level of development, demonstrating that low-income countries are associated with low economic growth and, vice versa, demonstrates that economic progress has been made in high-



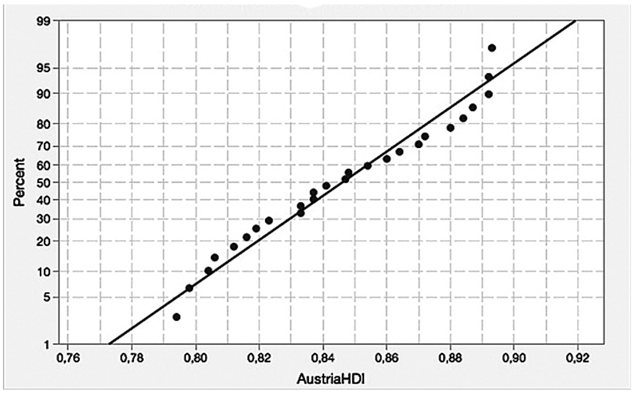
A1(a) Scatterplot of Austria GDP per capita ($) vs Austria HDI



A1(b) Normal Probability Plot of Austria GDP per capita ($); Austria HDI



A1(c) Normal Probability Plot of Austria GDP per capita ($)



A1(d) Normal Probability Plot of Austria HDI

Figure A1. Descriptive plots of the data (Austria).

Notes:

(a) A sample scatterplot of GDP per capita against the values of HDI based on the data extracted from World Bank (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=AT) and on Reports of United Nations (http://hdr.undp.org/en/data) over time for the period covering 27 years from 1990 to a year of 2017.

(b) A sample Normal Probability Plot of GDP per capita and HDI values over the same reporting years (1990–2017) based on the data of GDP per capita for Austria extracted from World Bank dataset (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=AT) and on the data of HDI for Austria extracted from Human Developments Reports of United Nations (http://hdr.undp.org/en/data).

(c) and (d) Separate Normal Probability plots of GDP per capita and HDI values over a twenty-five year period (1990–2017) (the same data). At the same time, the assumption that non-parametric data of Austria's GDP per capita still provides clear, reliable and relevant descriptive statistics.

income countries, with the high amounts of GDP per capita. Again, this reflects the complexity of permanently patterns of inequality. To conclude, we can see from simply comparing some of the measures of inequality that there are vast differences with regard to income, development and education levels across the world.

Analysis and Results

When looking at the relationship between two quantitative variables, researchers often hope to find a simple linear relationship. Sometimes the data reveals no relationship or other times a negative relationship. In this essay, let's look at some socioeconomic data of a country. The initial research question with this data is whether the relationship between an economic growth and how the levels of inequality affect it. The data covers 30 selected countries among three groups with regard to the level of development, and the variables are Human Development Index (HDI) values, Gross National Income (GNI) per capita, Education Levels measured by school enrollment ratios and GDP per capita measured in real terms per person. We have divided the three groups of countries by the levels of development i.e., developed countries (1st group), developing countries (2nd group) and countries of the third

Table 3

Command output from a simple regression of GDP per capita on HDI values for Austria for the period from 1990 to 2017

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
4536.05269	83.29%	82.59%	81.01%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-231253	24397	-9.48	< .0001	
AustriaHDI	315153	28819	10,94	< .0001	1

Regression Equation

AustriaGDPpercap($) = −231253 + 315153 AustriaHDI

Note. GDP per capita figures extracted from the World Bank (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=AT) and the values of HDI based on the data taken from the Reports of United Nations (http://hdr.undp.org/en/data) over time for the period covering 27 years from 1990 to a year of 2017.

world (3[d] group). The examples below will use the outputs from Minitab because it is simple and self-documenting.

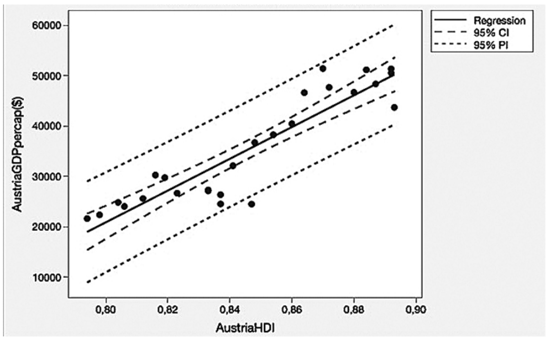
The following examples (with data from World Bank and United Nations Human Development Reports) represented the analysis of three selected countries, Austria, Mexico and Mozambique, taken from all three groups of countries, in which there are thirty nations in total and covering a twenty-seven-year period from 1990 to 2017. Despite the fact that a variety of countries is taken, the reason we present only one per each group is that each category reveals similar patterns and it would be unnecessary to characterize all thirty nations. Thus, instead of it we have selected countries with typical patterns of evaluating history inherent in each group.

Descriptive statistics of Austria

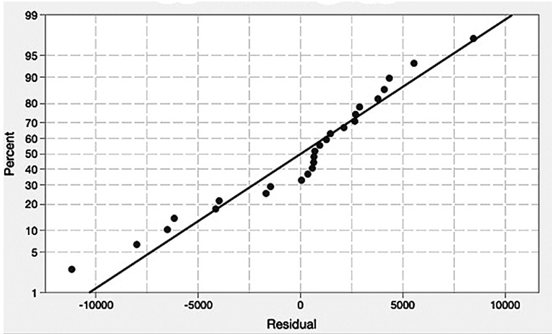
The scatterplot in figure A1(a) demonstrates that there seems to be a strong positive relationship between these two variables. One way to measure the strength of the relationship is correlation coefficient. For this reason, a Pearson correlation coefficient of Austria's GDP per capita measured in US dollars and HDI values is used where we test whether both variables are correlated. The model exhibits significant correlation of the linear regression model (Regression model: Pearson's $\rho = .9126$, $p < .0001$, $N = 25$). However, after running additional tests on normality, Anderson-Darling tests, the data revealed that one variable, namely, GDP per capita do not follow normal distribution (AD-value: 1.33, p = .005). This can be visually seen in Figures A1(c) and A1(d). Thus, new Spearman's rank correlation coefficient analysis has been carried out, which is a nonparametric measure of rank correlation. Spearman's ρ indicates that GDP per capita is significantly correlated with inequality-adjusted HDI values ($\rho = .8733$, $p < .0001$, $N = 25$).

Before proceeding to the practical interpretation of the regression coefficients for the linear relationship of GDP per capita and inequality-adjusted HDI values, we would like to remind that regression coefficients present the mean change in the response variable for one unit of change in the predictor variable, meanwhile, holding other predictors in the model constant. Thus, we would like to illustrate this in the scatterplot with a fitted line below, where a Pearson's GDP per capita is used to model their HDI values. First, it is important to consider Minitab's session window output below. The scatterplot with a fitted line in Figure A2(a) illustrates the same regression results graphically.

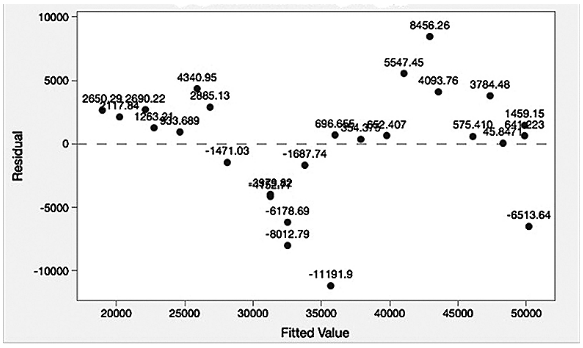
In this case regression equation is the following. The equation represents that the coefficient for HDI index is 315.153 in US dollars. The coefficient shows that for every additional index can expect GDP per capita to rise by an average of 315.153 US dollars. The R-squired is a statistical measure which tells how close the data are to the fitted regression line. Table 3 demonstrates



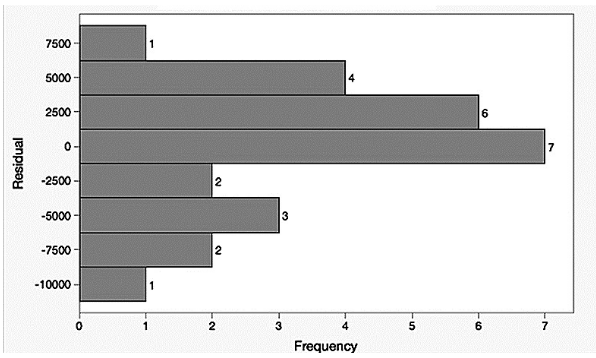
A2(a) Fitted Line Plot for Linear Model



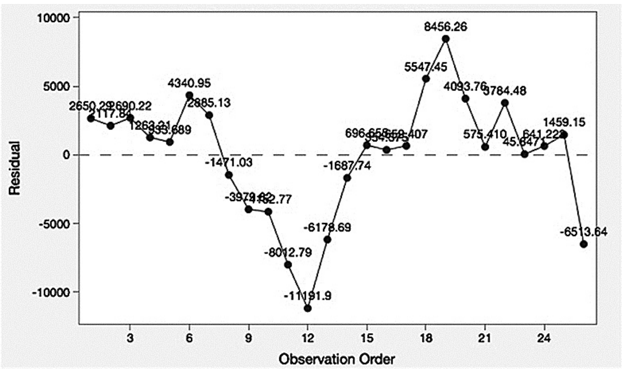
A2(b) Normal Probability Plot (response is Austria GDP per capita ($))



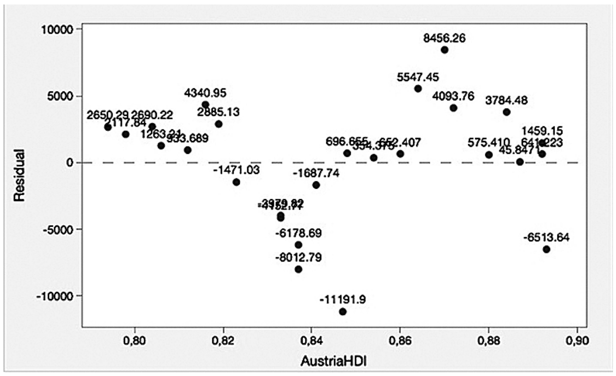
A2(c) Versus Fits (response is Austria GDP per capita ($))



A2(d) Histogram (response is Austria GDP per capita ($))



A2(e) Versus Order (response is Austria GDP per capita ($))



A2(f) Residuals versus Austria HDI (response is Austria GDP per capita ($))

Figure A2. Descriptive statistics of a simple liner regression of Austria.

Notes:

(a) A sample scatterplot with a fitted line plot of GDP per capita against the values of HDI based on the data extracted from World Bank (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=AT) and on Reports of United Nations (http://hdr.undp.org/en/data) over time for the period covering 27 years from 1990 to a year of 2017. At the same time, the 95% confidence and prediction intervals are also displayed.

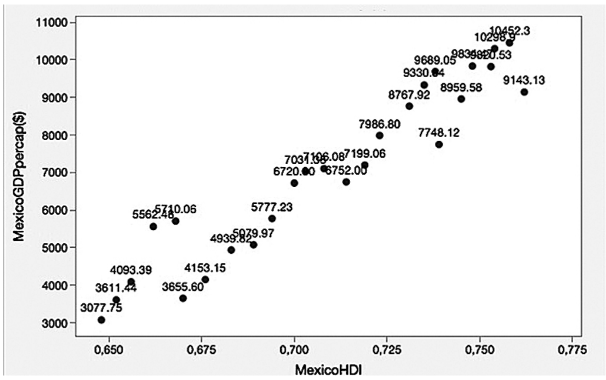
(b) A sample Normal Probability Plot of Residuals with GDP per capita as a response over the same reporting years (1990–2017) based on the data of GDP per capita for Austria extracted from World Bank dataset (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=AT) and on the data of HDI for Austria extracted from Human Developments Reports of United Nations (http://hdr.undp.org/en/data).

(c) Residual plots versus fits with GDP per capita as a response over a twenty-five-year period (1990–2017) (the same data).

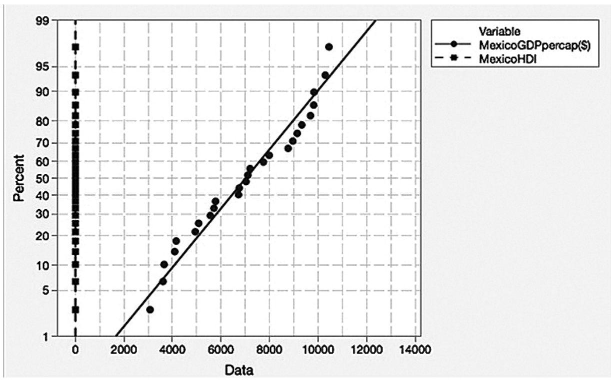
(d) A sample histogram of residuals with GDP per capita as a response over the same reporting years (1990–2017) based on the same data.

(e) A plot of residuals versus order with GDP per capita as a response over the same reporting years (1990–2017) based on the same data.

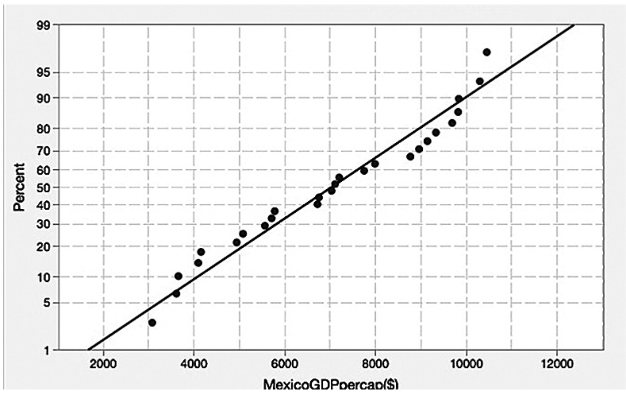
(f) A plot of residuals versus a separate variable of HDI values. At the same time, the assumption that non-parametric data of Austria's GDP per capita still provides clear, reliable and relevant descriptive statistics.



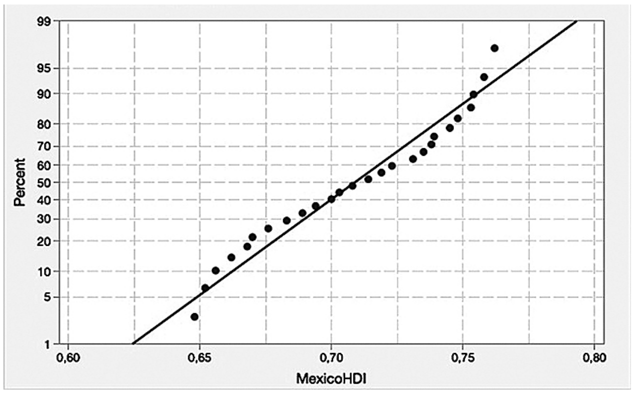
A3(a) Scatterplot of Mexico GDP per cap ($) vs Mexico HDI



A3(b) Normal Probability Plot of Mexico GDP per cap ($); Mexico HDI



A3(c) Normal Probability Plot of Mexico GDP per cap ($)



A3(d) Normal Probability Plot of Mexico HDI

Figure A3. Descriptive plots of the data (Mexico).

Notes:

(a) A sample scatterplot of GDP per capita against the values of HDI based on the data extracted from World Bank (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=MX) and on Reports of United Nations (http://hdr.undp.org/en/data) over time for the period covering 27 years from 1990 to a year of 2017.

(b) A sample Normal Probability Plot of GDP per capita and HDI values over the same reporting years (1990–2017) based on the data of GDP per capita for Mexico extracted from World Bank dataset (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=MX) and on the data of HDI for Mexico extracted from Human Developments Reports of United Nations (http://hdr.undp.org/en/data).

(c) and (d) Separate Normal Probability plots of GDP per capita and HDI values over a twenty-five-year period (1990–2017) (the same data).

that the regression model accounts for 82.59% of the variance.

The presented fitted line in figure A2(a) graphically illustrates the same information. If we move right or left along the x-axis by an amount that represents a one unit change in HDI, the fitted line decreases or increases by 315.153 US dollars. However, these HDIs are for developed countries and range from 0.794 to 0.893. The relationship is only valid within this data range, so we would not actually shift upward or downward along the line by a full unit of index in this case.

If the fitted line was flat (a slope coefficient of zero), the expected value for GDP per capita would stay unchanged no matter how far we go upward or downward the line. Thus, a very small p-value suggests that the slope is not equal to zero, which subsequently, indicates that changes in the predictor variable are associated with changes in the response variable.

The reason we used a fitted line plot is that it brings math to life. Nevertheless, fitted line scatterplots may only display the results from simple regression, that is to say one predictor variable and the response.

Descriptive statistics of Mexico

The scatterplot in figure A3(a) demonstrates that there seems to be a strong positive relationship between these two variables. One

Table 4
Command output from a simple regression of GDP per capita on HDI values for Mexico for the period from 1990 to 2017

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
694.104223	91.27%	90.91%	89.60%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-35946	2715	-13.24	< .0001	
MexicoHDI	606620	3826	15.84	< .0001	1

Regression Equation

MexicoGDPpercap($) = -35946 + 606620 MexicoHDI

Note. GDP per capita figures extracted from the World Bank (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=MX) and the values of HDI based on the data taken from the Reports of United Nations (http://hdr.undp.org/en/data) over time for the period covering 27 years from 1990 to a year of 2017.

way to measure the strength of the relationship is with correlation coefficient. For this reason, we have run a Pearson correlation coefficient of Mexico's GDP per capita measured in US dollars and HDI values where it is tested whether both variables are correlated with each other. The model exhibits significant correlation of the linear regression model (Regression model: Pearson's $\rho = 0.9553$, $p < .0001$, $N = 25$). Importantly, after running additional tests on normality, Anderson-Darling Tests, the data revealed that both variables, follow normal distribution (GDP per capita: AD-value: 0.43, $p = .2895$; HDI: AD-value: 0.44, $p = .2623$). This can be visually seen in figures A3(c) and A3(d).

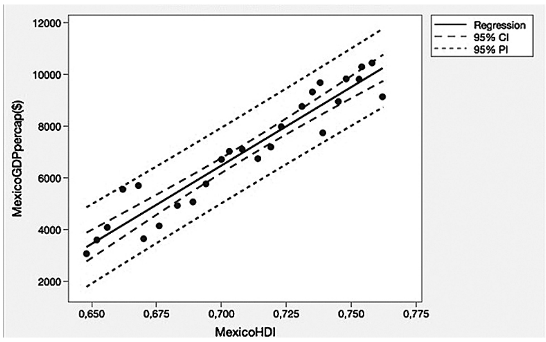
Before proceeding to the practical interpretation of the regression coefficients for the linear relationship of GDP per capita and inequality-adjusted HDI values, we would like to remind that regression coefficients present the mean change in the response variable for one unit of change in the predictor variable, meanwhile, holding other predictors in the model constant. Thus, let's illustrate this in the scatterplot with a fitted line below, where a Pearson's GDP per capita is used to model their HDI values. First, it is important to consider Minitab's session window output below. The scatterplot with a fitted line in figure A4(a) illustrates the same regression results graphically.

In this case regression equation is the following. The equation represents that the coefficient for HDI index is 606.620 in US dollars. The coefficient shows that for every additional index figure we can expect GDP per capita to rise by an average of 606.620 US dollars. The R-squired is a statistical measure which tells how close the data is to the fitted regression line. Table 4 demonstrates that the regression model accounts for 90.91% of the variance.

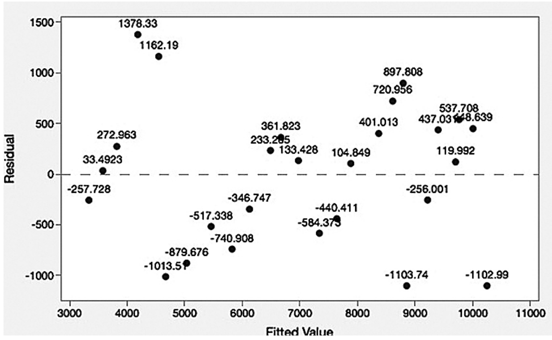
The fitted line in figure A4(a) graphically illustrates the same information. If we move right or left along the x-axis by an amount that represents a one unit change in HDI, the fitted line decreases or increases by 606.620 US dollars. However, these HDIs are for developed countries and range from 0.648 to 0.762. The relationship is only valid within this data range, so we would not actually shift upward or downward along the line by a full unit of index in this case.

If the fitted line was flat (a slope coefficient of zero), the expected value for GDP per capita would stay unchanged no matter how far you go upward or downward the line. Thus, a very small p-value suggests that the slope is not equal to zero, which subsequently, indicates that changes in the predictor variable are associated with changes in the response variable.

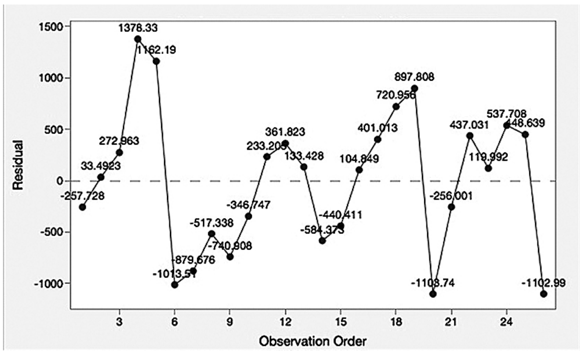
The reason we used a fitted line plot is that it brings the math to life. Nevertheless, fitted line scatterplots may only display the results from



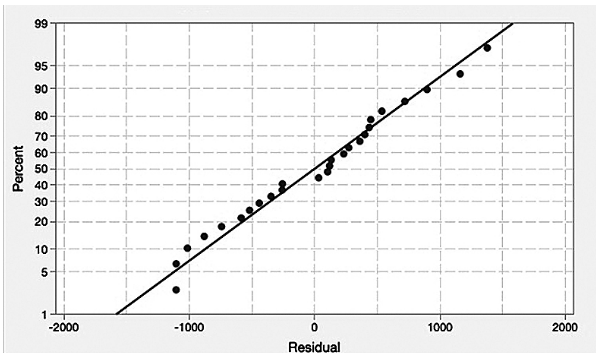
A4(a) Fitted Line Plot
for Linear Model



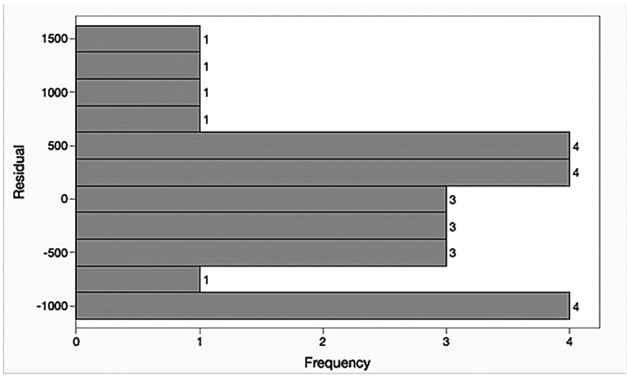
A4(b) Versus Fits
(response is Mexico GDP per capita ($))



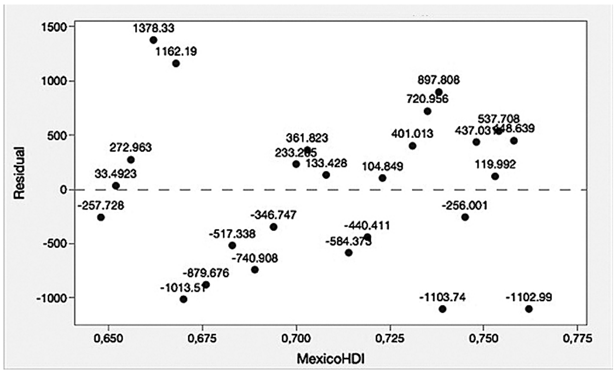
A4(c) Versus Order
(response is Mexico GDP per capita ($))



A4(d) Normal Probability Plot (response is Mexico GDP per capita ($))



A4(e) Histogram
(response is Mexico GDP per capita ($))



A4(f) Residuals versus Mexico HDI (response is Mexico GDP per capita ($))

Figure A4. Descriptive statistics of a simple liner regression of Mexico.

Notes.

(a) A sample scatterplot with a fitted line plot of GDP per capita against the values of HDI based on the data extracted from World Bank (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=MX) and on Reports of United Nations (http://hdr.undp.org/en/data**)** over time for the period covering 27 years from 1990 to a year of 2017. At the same time, the 95% confidence and prediction intervals are also displayed.

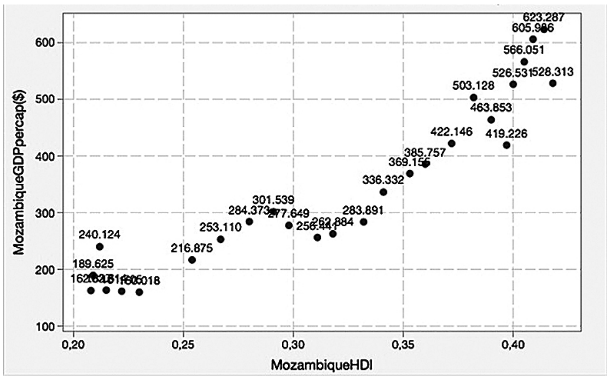
(b) A sample Normal Probability Plot of Residuals with GDP per capita as a response over the same reporting years (1990–2017) based on the data of GDP per capita for Mexico extracted from World Bank dataset (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=MX) and on the data of HDI for Mexico extracted from Human Developments Reports of United Nations (http://hdr.undp.org/en/data**)**.

(c) Residual plots versus fits with GDP per capita as a response over a twenty-five-year period (1990–2017) (the same data).

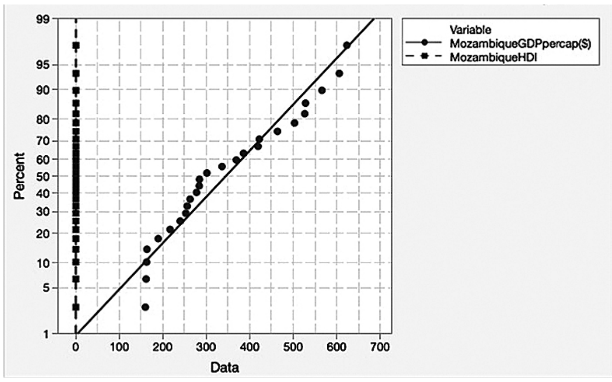
(d) A sample histogram of residuals with GDP per capita as a response over the same reporting years (1990–2017) based on the same data.

(e) A plot of residuals versus order with GDP per capita as a response over the same reporting years (1990–2017) based on the same data.

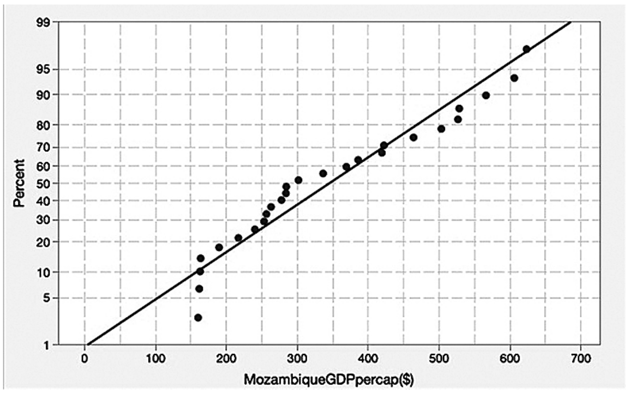
(f) A plot of residuals versus a separate variable of HDI values.



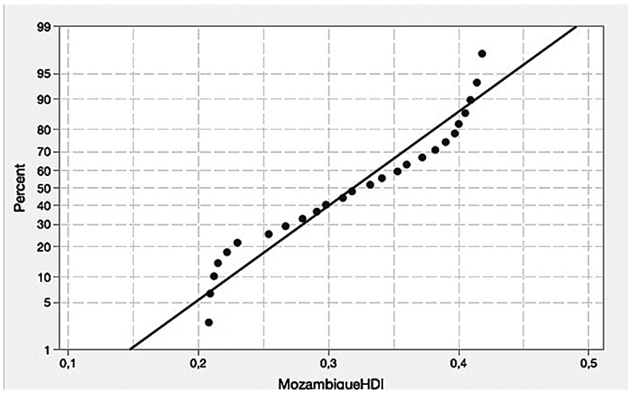
A5(a) Scatterplot of Mozambique GDP per cap ($) vs Mozambique HDI



A5(b) Normal Probability Plot of Mozambique GDP per cap ($); Mozambique HDI



A5(c) Normal Probability Plot of Mozambique GDP per cap ($)



A5(d) Normal Probability Plot of Mozambique HDI

Figure A5. Descriptive plots of the data (Mozambique).

Notes:

(a) A sample scatterplot of GDP per capita against the values of HDI based on the data extracted from World Bank (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=MZ) and on Reports of United Nations (http://hdr.undp.org/en/data) over time for the period covering 27 years from 1990 to a year of 2017.

(b) A sample Normal Probability Plot of GDP per capita and HDI values over the same reporting years (1990–2017) based on the data of GDP per capita for Mozambique extracted from World Bank dataset (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=MZ) and on the data of HDI for Mozambique extracted from Human Developments Reports of United Nations (http://hdr.undp.org/en/data).

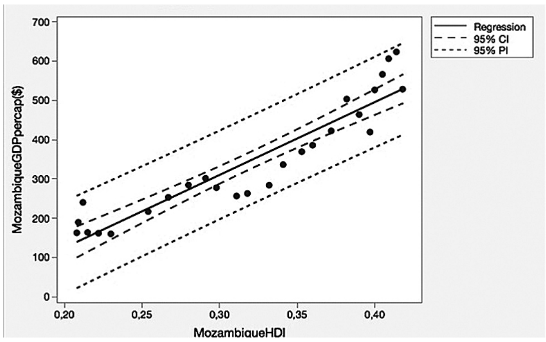
(c) and (d) Separate Normal Probability plots of GDP per capita and HDI values over a twenty-five year period (1990–2017) (the same data).

simple regression, that is to say one predictor variable and the response.

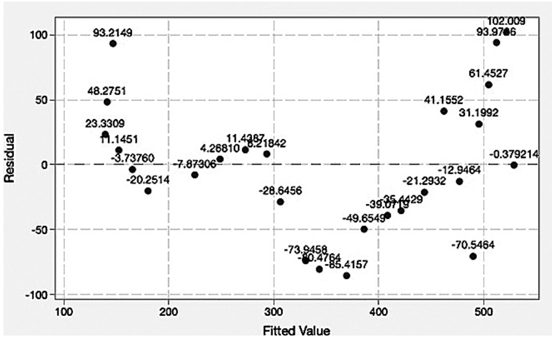
Descriptive statistics of Mozambique

The scatterplot in figure A5(a) also demonstrates that there exists a strong positive relationship between these two variables. One way to measure the strength of the relationship is correlation coefficient. For this reason, we have run a Pearson correlation coefficient of Mozambique's GDP per capita measured in US dollars and HDI values where we test whether both variables are correlated. The model exhibits significant correlation of the linear regression model (Regression model: Pearson's $\rho = 0.9334$, $p < .0001$, $N = 25$). Importantly, after running additional tests on normality, Anderson-Darling tests, the data revealed that both variables, follow normal distribution (GDP per capita: AD-value: 0.63, $p = .0900$; HDI: AD-value: 0.68, $p = .0666$). This can be visually seen in figures A5(c) and A5(d).

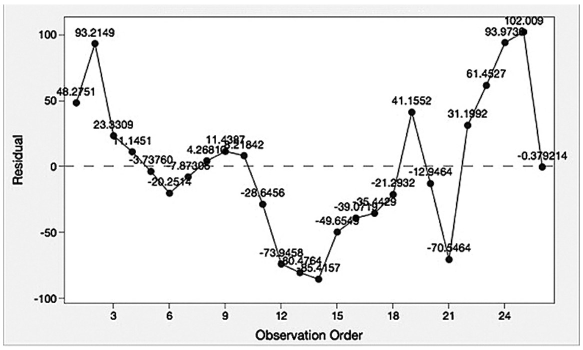
Thus, as usual, before proceeding to the practical interpretation of the regression coefficients for the linear relationship of GDP per capita and inequality-adjusted HDI values, we would like to illustrate this in the scatterplot with a fitted line below, where we are going to use a Pearson's GDP per capita to model their HDI values. First, it is important to consider



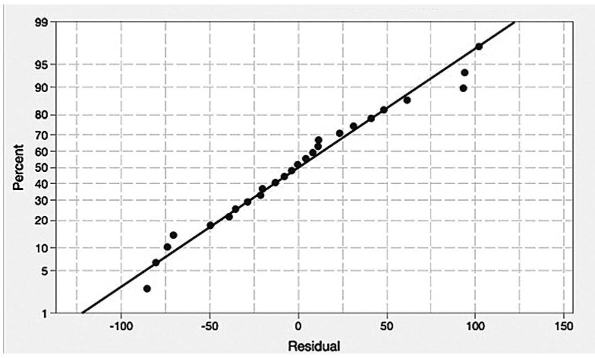
A6(a) Fitted Line Plot
for Linear Model



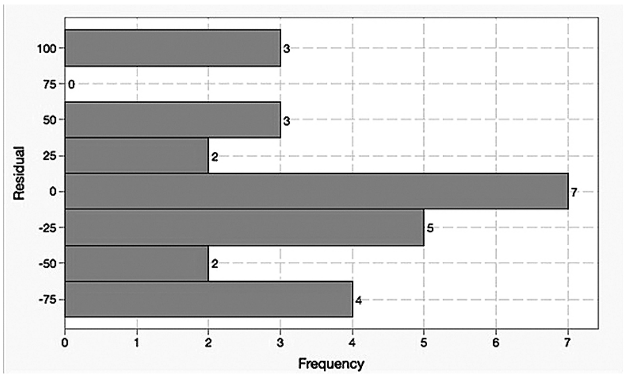
A6(b) Versus Fits (response is Mozambique GDP per cap ($))



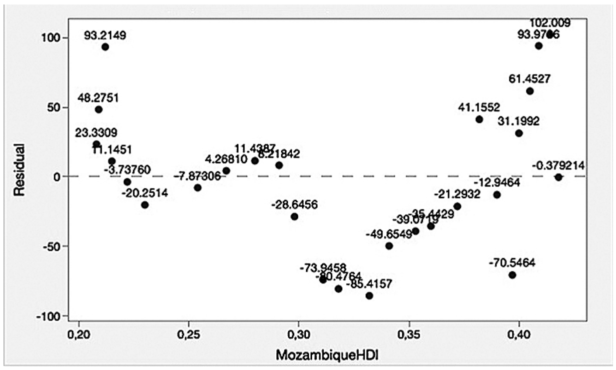
A6(c) Versus Order (response is Mozambique GDP per cap ($))



A6(d) Normal Probability Plot (response is Mozambique GDP per cap ($))



A6(e) Histogram
(response is Mozambique GDP per cap ($))



A6(f) Residuals versus Mozambique HDI
(response is Mozambique GDP per cap ($))

Figure A6. Descriptive statistics of a simple liner regression of Mozambique.

Notes:

(a) A sample scatterplot with a fitted line plot of GDP per capita against the values of HDI based on the data extracted from World Bank (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=MZ) and on Reports of United Nations (http://hdr.undp.org/en/data**)** over time for the period covering 27 years from 1990 to a year of 2017. At the same time, the 95% confidence and prediction intervals are also displayed.

(b) A sample Normal Probability Plot of Residuals with GDP per capita as a response over the same reporting years (1990–2017) based on the data of GDP per capita for Mozambique extracted from World Bank dataset (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=MZ) and on the data of HDI for Mozambique extracted from Human Developments Reports of United Nations (http://hdr.undp.org/en/data**).**

(c) Residual plots versus fits with GDP per capita as a response over a twenty-five year period (1990–2017) (the same data).

(d) A sample histogram of residuals with GDP per capita as a response over the same reporting years (1990–2017) based on the same data.

(e) A plot of residuals versus order with GDP per capita as a response over the same reporting years (1990–2017) based on the same data.

(f) A plot of residuals versus a separate variable of HDI values.

Table 5

Command output from a simple regression of GDP per capita on HDI values for Mexico based on a period from 1990 to 2017

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
53.6106047	87.14%	86.60%	84.76%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-245.99	47.51	-5.18	< .0001	
MozambiqueHDI	1853.3	145.4	12.75	< .0001	1

Regression Equation

MozambiqueGDPpercap($) = -245.99 + 1853.3 MozambiqueHDI

Note. GDP per capita figures extracted from the World Bank (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=MZ) and the values of HDI based on the data taken from the Reports of United Nations (http://hdr.undp.org/en/data) over time for the period covering 27 years from 1990 to a year of 2017.

Minitab's session window output below. The scatterplot with a fitted line in figure A6(a) illustrates the same regression results graphically.

In this case regression equation is the following. The equation represents that the coefficient for HDI index is 1,853.3 in US dollars. The coefficient shows that for every additional index figure we can expect GDP per capita to rise by an average of 1,853.3 US dollars. The R-squired is a statistical measure which tells how close the data is to the fitted regression line. Table 5 demonstrates that the regression model accounts for 84.76% of the variance.

The fitted line in figure A6(a) graphically illustrates the same information. If we move right or left along the x-axis by an amount that represents a one unit change in HDI, the fitted line decreases or increases by 1853.3 US dollars. However, these HDIs are for developed countries and range from 0.209 to 0.418. The relationship is only valid within this data range, so we would not actually shift upward or downward along the line by a full unit of index in this case.

If the fitted line was flat (a slope coefficient of zero), the expected value for GDP per capita would stay unchanged no matter how far you go upward or downward the line. Thus, a very small p-value suggests that the slope is not equal to zero, which subsequently, indicates that changes in the predictor variable are associated with changes in the response variable.

Conclusions

After carrying out statistical analysis, let's confirm that inequality slows economic growth. Low-income countries associated with high levels of inequality tend to grow more slowly in economic terms. Our findings are supported by the similar results of Ortiz and Cummins (2011). Thus, economic growth is mostly associated with countries characterized by well-developed economic policy, which supports and promotes free trade. It can lead to economic growth and, subsequently, to poverty reduction.

In other words, increasing wealth is seen to decrease poverty. However, based on some relevant studies, there is a contradictory evidence in relation to this that economic growth does not automatically result in reduced levels of inequality; rather it is considered to be a main factor in enriching the rich and further impoverishing the poor.

According to the study by World Bank (World Development Report, 2017), inequality can slow economic growth and therefore be seen as negative by economists. Economic growth itself is unlikely to result in poverty reduction. From our opinion, it is incorrect to fully ignore economic growth.

We strongly believe, and support the opinion by analytical data, that economic growth is the main stimulator of the level of inequality. That is to say, if a nation is wealthy enough, the society in it will live in prosperity.

References

Firebaugh, Glenn. (2004). Does industrialization no longer benefit poor countries? A comment on Arrighi, Silver, and Brewer. *Studies in Comparative International Development, 39*(1), 99–105. Retrieved from http://sites.jmu.edu/benbrewer/files/2011/01/SCID-Firebaugh.pdf.

Human Development Data (1990–2015). *Human Development Reports*. United Nations Development Programme. Retrieved from http://hdr.undp.org/en/data.

Milanovic, Branko, Lindert, Peter, & Williamson, Jeffrey. (2007, October). Measuring Ancient Inequality. National Bureau of Economic Research. *MPRA Paper No. 5388*. Retrieved from https://mpra.ub.uni-muenchen.de/5388/1/MPRA_paper_5388.pdf

Milanovic, Branko, Lindert, Peter H., & Williamson, Jeffrey G. (2010). Pre-Industrial Inequality. *The Economic Journal, 121* (March), 255–272. doi:10.1111/j.1468–0297.2010.02403.x. Retrieved from https://gc.cuny.edu/CUNY_GC/media/CUNY-Graduate-Center/PDF/Centers/LIS/Milanovic/papers/2011/MLW_final.pdf.

Ortiz, Isabel, & Cummins, Matthew. (2011). *Global Inequality: Beyond the Bottom Billion. A Rapid Review of Income Distribution in 141 Countries*. United Nations Children's Fund (UNICEF). Retrieved from https://www.unicef.org/socialpolicy/files/Global_Inequality.pdf.

Sen, A. (1999). *Development as freedom*. Oxford, England: Oxford University Press. Retrieved from http://www.c3l.uni-oldenburg.de/cde/OMDE 625/Sen/Sen-intro.pdf.

World Bank Open Data. World Bank. Retrieved from https://data.worldbank.org.

World Development Report 2017. *Governance and the Law*. (2017). Washington, DC: The World Bank. Retrieved from http://www.worldbank.org/en/publication/wdr2017.

World Population Ageing 2009. (2009). ESA/P/WP/212. New York, NY: United Nations. Retrieved from http://www.un.org/esa/population/publications/WPA2009/WPA2009_WorkingPaper.pdf.

World Population Ageing Report 2015 [Report]. (2015). ST/ESA/SER.A/390. New York, NY: United Nations. Retrieved from http://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Report.

World Urbanization Prospects. The 2007 Revision. Highlights. (2008). New York, NY: United Nations. Retrieved from http://www.un.org/esa/population/publications/wup2007/2007WUP_Highlights_web.

Стимулирует ли социальное неравенство экономический рост?
(на примерах выбранных развивающихся стран)

Алина Пухаева[1], Елена Мирошина (Силантьева)[2]

В статье критически рассматривается концепция социального неравенства и предлагаются способы ее определения на фоне широкого спектра факторов, определяющих неравенство среди самых богатых и беднейших стран. В ней также содержатся обобщенные показатели неравенства между тремя группами стран путем сопоставления некоторых макроэкономических показателей социально-экономического неравенства. Затем мы проверили наличие линейной зависимости между двумя количественными переменными. Используя данные Всемирного банка и докладов Организации Объединенных Наций по человеческому развитию, мы провели анализ отдельных стран, взятых из трех групп стран (всего тридцать стран), за период с 1990 по 2017 год. После проведения статистического анализа мы доказали, что неравенство замедляет экономический рост.

Ключевые слова: неравенство; доход; ВВП на душу населения; бедность; HDI
JEL classification: I31, J11, J24, D63

[1] Международный финансовый факультет, Финансовый университет, Москва, Россия; alexandra.puhaeva@gmail.com

[2] Кандидат экономических наук, Департамент мировой экономики и мировых финансов, Финансовый университет, Москва, Россия; EAsila@yandex.ru

Review of Business and Economics Studies
2018, Vol. 6, No. 1, 56–77

DOI: 10.26794/2308-944X-2018-6-1-56–77

Country Risk in International Investment Its' structure and methods of estimation

Alexey Ivkin[1]

[1]Student of Financial University under the Government of the Russian Federation, master program «International Finanance», Moscow, Russia; vok-zal@mail.ru; +7 (916) 246–83–89.
Scientific Supervisor — A.I. Ilinsky, D. Sc. in engineering, professor of the Macroaconomic Regulation Faculty, Financial University under the Government of the Russian Federation, Moscow, Russia.

Abstract

In this paper, the issues of country risk assessment in economic security and sustainability context are investigated. The main object of research is country risk and its structural components. The scientific paper's main goal is to analyze valuation methods of country risk from different perspectives and suggest a model for country risk measurement that allows adequately evaluate country risk, economic security and economic sustainability level and dynamics, including structural components and their relationships. The paper approaches several main tasks. First, to highlight the importance of country risk evaluation and its assessment in growing global markets, analyzing causes and elements of country risk based on other scientific researches. Second, to explore and clarify advantages and disadvantages of the methods of country's risk assessment, as well as to investigate sources of country risk and ways how to manage the risk. Third, to apply quantitative and qualitative methods for analysis, formulate, create and present the model of country's risk assessment in economic security and sustainability context, which will identify factors, influencing country risk and determine their direct and indirect relationship between each other. The last task is to verify practical suitability of the model of country's risk assessment by performing empirical analysis all over the world, identifying directions for mitigating risk effects.

Keywords: country risk analysis; credit ratings; debt; structured qualitative methods; discriminant analysis; *k*-nearest neighbours (*k*-NN) classification; distance metrics; classification and regression trees (CART); support vector machine (SVM); discriminant analysis; missing data; Kolomogorov-Smirnov test; Anderson-Darling test; non-parametric methods.
JEL classification: C53, O16, O47

Each business operation causes some kind of risk. When business operations occur in international dimension, they bring additional risks, which are not typical for domestic operations. These additional risks are called country risks and usually include risks arising from a variety of national differences in policies, geography, economic structures, socio-political institutions, and currencies. Country risk analysis (CRA) tries to solve this problem by identifying the potential for these risks to decrease the expected return of cross-border investments.

Concept of "Country risk" began to be widely used in the 1970s. It was originally more professionally oriented in the sense that it aimed at addressing the concrete issue of a particular business in a particular country and was generally used by the banking industry.

Every year it becomes more and more difficult to analyze and predict changes in the financial, economic and political sectors of business. The importance of country risk analysis is now more understandable and potential for it is growing by establishing more and more country risk rating agencies, which combine a wide range of qualitative and quantitative information regarding alternative measures of economic, financial, and political risk into associated composite risk ratings. However, the accuracy of any rating agency with regard to any or all of these measures is open to question.

Globalization, after undermining the old definition of economic security, is found at the centre of a new definition that emphasizes the risks of unexpected shocks and economic volatility. The new definition must capture the causal consequences of globalization accurately and establish explicit benchmarks for assessing globalization's effects on economic security and country's economic sustainability.

Country risk assessment has been analyzed by different authors but in quite narrow way, in this paper the concept of country risk and influencing factors are presented in an extended view.

Following new scientific novelties in economics were discovered:

Expanded and consolidated overview of analyzes of country risk concept, its components and arising problems were analyzed in another angle which allowed to identify new possibilities and challenges for creating new model for assessment of country risk.

Broader analysis of country risk—includes not only political risk, but as well socio-economical aspects, presents clear and analyzed new concept which was not assumed in previous researches.

There are many studies related to country risk, its financial integration in a country, the impact on economics, and other aspects of country's welfare. To summarize the analysis of scientific literature about country risk, it is obvious that researchers are analyzing country risk approach only partially, not adapting the concept to growing globalization topic, which definitely makes changes in country risk approach. Country risk concept should be analyzed and understand in a broader way. This updated approach I will discuss further.

Methodologies of Country Risk and Methods for Its Valuation

Credit Ratings

Credit rating agencies (CRAs) play a key role in financial markets by helping to reduce the informative asymmetry between lenders and investors, on one side, and issuers on the other side, about the creditworthiness of companies or countries. CRAs' role has expanded with financial globalization and has received an additional boost from Basel II that incorporates the ratings of CRAs into the rules for setting weights for credit risk. Ratings tend to be sticky, lagging markets, and overreact when they do change. This overreaction may have aggravated financial crises in the recent past, contributing to financial instability and cross-country contagion.

A credit rating is a current opinion and measure of the risk of an obligor with respect to a specific financial obligation based on all available information. For this purpose, S&P and Fitch define risk as the probability of default (PD), whereas Moody's define it as 'loss'.

The logic underlying the existence of CRAs is to solve the problem of the informative asymmetry between lenders and borrowers regarding the creditworthiness of the latter. Issuers with lower credit ratings pay higher interest rates embodying larger risk premiums than higher rated issuers. Moreover, ratings determine the eligibility of debt and other financial instruments for the portfolios of certain institutional investors due to national regulations that restrict investment in speculative-grade bonds.

Standard and Poor's ratings seek to capture only the forward-looking probability of the occurrence of default. They provide no assessment of the expected time of default or mode of default resolution and recovery values.

By contrast, Moody's ratings focus on the Expected Loss (EL) that is a function of both Probability of Default (PD) and the expected Recovery Rate (RE). Thus $EL = PD (1 — RE)$.

Fitch's ratings also focus on both PD and RE. They have a more explicitly hybrid character in that analysts are also reminded to be forward-looking and to be alert to possible discontinuities between past track records and future trends.

Models Used

A variety of statistical methods was employed to estimate models of debt rescheduling in the studies cited above. Since most authors chose their dependent variable to be a discrete binary variable, which took on the value one when a country 'rescheduled' within a given time-period and zero otherwise, the statistical methods used have been those designed for dichotomous dependent variables. These methods include discriminant analysis, linear-probability, probit, and logit models. In this section, we briefly describe each of these methods and then discuss criteria to use when choosing among them. Perforce, our discussion will be brief.

The methods used by the banks and other agencies for country risk analysis can broadly be classified as qualitative or quantitative. However, many agencies amalgamate both qualitative and quantitative information into a single index or rating. The data was collected from various sources that include expert panel, survey, staff analysis, and published data sources. The country risk index could be either ordinal or scalar. A survey conducted by the US Export-Import Bank in 1976 categorized various methods of country risk appraisal used mainly by the banks into one of four types: (1) full qualitative method, (2) structured qualitative method, (3) checklist method, and (4) other quantitative method. Since our focus in this paper is on quantitative methods, we will only briefly discuss the other three categories.

Discriminant analysis. Discriminant analysis finds a set of prediction equations based on independent variables that are used to classify individuals into groups. There are two possible objectives in a discriminant analysis: finding a predictive equation for classifying new individuals or interpreting the predictive equation to better understand the relationships that may exist among the variables.

Objectives. The main objectives of discriminant analisys are:

Development of discriminant functions.

Examination of whether significant differences exist among the groups, in terms of the predictor variables.

Determination of which predictor variables contribute to most of the intergroup differences.

Evaluation of the accuracy of classification.

Discriminant analysis linear equation. A involves the determination of a linear equation like regression that will predict which group the case belongs to. The form of the equation or function is:

$$D = v_1X_1 + v_2X_2 + v_3X_3 + \dots + v_iX_i + a,$$

where

D = discriminate function

v = the discriminant coefficient or weight for that variable

X = respondent's score for that variable

a = a constant

i = the number of predictor variables.

Assumptions of discriminant analysis. The main oassumptions of discriminant analisys are:

The observations are a random sample.

Each predictor variable is normally distributed.

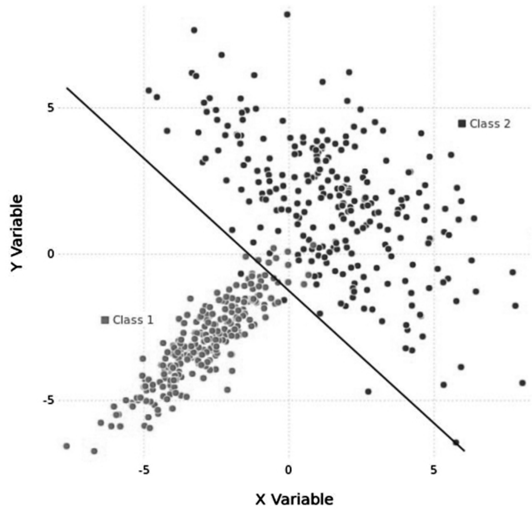


Figure 1. Example of linear discriminant analisys.

Each of the allocations for the dependent categories in the initial classification are correctly classified.

There must be at least two groups or categories, with each case belonging to only one group so that the groups are mutually exclusive and collectively exhaustive (all cases can be placed in a group).

Each group or category must be well defined, clearly differentiated from any other group(s) and natural.

The groups or categories should be defined before collecting the data.

The attribute(s) used to separate the groups should discriminate quite clearly between the groups so That group or category overlap is clearly non-existent or minimal.

Group sizes of the dependent should not be grossly different and should be at least five times the number of independent variables.

K-nearest neighbours (*k*-NN) algorithm. The *k*-NN (k-nearest neighbours) algorithm is a classification algorithm that can apply to question classification. However, its time complexity will increase linearly with the increase of training set size, which constrains the actual application effects of this algorithm. Simply stated, *k*-NN is an algorithm that classifies the new cases based on similarity measures or distance measures of pair of observations such as euclidean, cosine, etc. *k*-NN algorithm is a lazy learner i.e. it does not learn anything from the training tuples and simply uses the training data itself for classification. It is a non-parametric method used for classification and regression.

Different types of prediction using data mining techniques are:

(1) *Classification:* predicting into what category or class a case falls.

(2) *Regression*: predicting what number value a variable will have (if a variable varies with time, it is called 'time series' prediction).

Classification problems aim to identify the characteristics that indicate the group to which each case belongs. This pattern can be used both to understand the existing data and to predict how new instances will behave. Data mining creates classification models by examining already classified data (cases) and inductively finding a predictive pattern.

k-NN for classification. In pattern recognition, the *k*-NN algorithm is a method for classifying objects based on closest training examples in the feature space. *k*-NN is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification.

Figure 2 shows the *k*-NN decision rule for $K = 1$ and $K = 4$ for a set of samples divided into 2 classes. In Figure 2(a), an unknown sample is classified by using only one known sample; in Figure 2(b) more than one known sample is used. In the last case, the parameter K is set to 4, so that the closest four samples are considered for classifying the unknown one. Three of them belong to the same class, whereas only one belongs to the other class. In both cases, the unknown sample is classified as belonging to the class on the left.

Distance Metric. As mentioned before KNN makes predictions based on the outcome of the K neighbors closest to that point. Therefore, to make predictions with KNN, we need to define a metric for measuring the distance between the query point and cases from the examples sample. One of the most popular choices to measure this distance is known as Euclidean. Other measures include Euclidean squared, City-block, and Chebychev.

$$D(x,p) = \begin{cases} \sqrt{(x-p)^2} \\ (x-p)^2 \\ |x-p| \\ Max(|x-p|), \end{cases}$$

where x and p are the query point and a case from the examples sample, respectively.

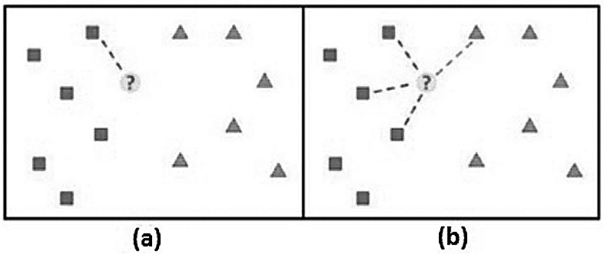


Figure 2. (a) The 1-NN decision rule: the point is assigned to the class on the left; (b) the KNN decision rule, with K = 4 the point is assigned to the class on the left as well.

The advantages of the *k*-NN method are as follows:

1. Analytically tractable.
2. Simple implementation.
3. Nearly optimal in the large sample limit ($NN \to \infty$).
4. Uses local information, which can yield highly adaptive behaviour.
5. Lends itself very easily to parallel implementations.

The drawbacks of the k-NN method are as follows:

1. *k*-NN algorithm is that it is a *lazy learner*, i.e. it simply uses the training data itself for classification.
2. Result of this is that the method does not learn anything from the training data, which can result in the algorithm not generalizing well. Further, changing K can change the resulting predicted class label.
3. Also algorithm may not be robust to noisy data.
4. To predict the label of a new instance the kNN algorithm will find the K closest neighbours to the new instance from the training data, the predicted class label will then be set as the most common label among the K closest neighbouring points.
5. The algorithm needs to compute the distance and sort all the cases at each prediction, which can be slow if there are a large number of training examples.
6. Large storage requirements.
7. Computationally intensive recall.
8. Highly susceptible to the curse of dimensionality.

Classification and regression trees (CART) method. CART is one method of machine learning which the exploration method is done by decision tree technique. The method developed by Leo Breiman, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone is classification technique by using binary recursive partition algorithm.

Classification tree methods such as CART are convenient way to produce a prediction rule from a set of observations described in terms of a vector of features and a response value. The aim is to define a general prediction rule, which can be used to assign a response value to the cases solely on the bases of their predictor (explanatory) variables. Tree-structured classifications are not based on assumptions of normality and user-specified model statements, as are some conventional methods such as discriminant analysis and ordinary least square regression.

Tree based decision methods are statistical systems that mine data to predict or classify future observations based on a set of decision rules and are sometimes called rule induction methods because the reasoning process behind them is clearly evident when browsing the trees.

In CART, the observations are successively separated into two subsets based on associated variables significantly related to the response variable; this approach has an advantage of providing easily comprehensible decision strategies.

CART can be applied either as a classification tree or as a regressive tree depending on whether the response variable is categorical or continuous. Tree based methods are not based on any stringent assumptions. These methods can handle large number of variables, are resistant to outliers, non-parametric, more versatile, can handle categorical variables, though computationally more intensive.

CART methodology. For building decision trees, CART uses so-called learning set- a set of historical data with pre-assigned classes for all observations. An algorithm known as recursive partitioning is the key to the nonparametric statistical method of CART. It is a step-by-step process by which a decision tree is constructed by either splitting or not splitting each node on the tree into two daughter nodes.

An attractive feature of the CART methodology is that because the algorithm asks a sequence of hierarchical questions, it is relatively simple to understand and interpret the results. The unique starting point of a classification tree is called a root node and consists of the entire learning set L at the top of the tree.

Steps in Cart. CART analysis consists of four basic steps. The first step consists of tree building, during which a tree is built using recursive splitting of nodes. Each resulting node is assigned a predicted class, based on the distribution of classes in the learning dataset which would occur in that node and the decision cost matrix. The assignment of a predicted class to each node occurs whether or not that node is subsequently split into child nodes.

The second step consists of stopping the tree building process. At this point a "maximal" tree has been produced, which probably greatly overfits the information contained within the learning dataset.

The third step consists of tree "pruning," which results in the creation of a sequence of simpler and simpler trees, through the cutting off of increasingly important nodes.

The fourth step consists of optimal tree selection, during which the tree which fits the information in the learning dataset, but does not overfit the information, is selected from among the sequence of pruned trees. Each of these steps will be discussed in more detail below.

The tree-building process starts by partitioning a sample or the root node into binary nodes based upon a very simple question of the form is $X \leqslant d$?, where X is a variable in the data set and d is a real number. Initially, all observations are placed in the root node. This node is impure or heterogenous because it contains observations of mixed classes. The goal is to devise a rule that will break up these observations and create groups or binary nodes that are internally more homogenous than the root node. Starting from the root node, and using, for example, the Gini diversity index as a splitting rule, the tree building process is as follows:

1. CART splits the first variable at all of its possible split points (at all of the values the variable assumes in the sample). At each possible split point of a variable, the sample splits into binary or two child nodes. Cases with a "yes" response to the question posed are sent to the left node and those with "no" responses are sent to the right node.

2. CART then applies its goodness-of-split criteria to each split point and evaluates the reduction in impurity that is achieved using the formula

$$\Delta i(s, t) = i(t) - p_L[i(t_L)] - p_R[i(t_R)],$$

3. CART selects the best split of the variable as that split for which the reduction in impurity is highest.

4. Steps 1–3 are repeated for each of the remaining variables at the root node.

5. CART then ranks all of the best splits on each variable according to the reduction in impurity achieved by each split.

6. It selects the variable and its split point that most reduced the impurity of the root or parent node.

7. CART then assigns classes to these nodes according to the rule that minimizes misclassification costs. CART has a built-in algorithm that takes into account user-defined variable misclassification costs during the splitting process. The default is unit or equal misclassification costs.

8. Because the CART procedure is recursive, steps 1–7 are repeatedly applied to each nonterminal child node at each successive stage.

9. CART continues the splitting process and builds a large tree. The largest tree is built if the splitting process continues until every observation constitutes a term in al node. Obviously, such a tree will have a large number of terminal nodes, which will be either pure or have very few cases (less than 10).

Summing up CART's strengths and weaknesses:

1. CART makes no distributional assumptions of any kind for dependent and independent variables. No variable in CART is assumed to follow any kind of statistical distribution.

2. The explanatory variables in CART can be a mixture of categorical and continuous.

3. CART has a built-in algorithm to deal with the missing values of a variable for a case, except when a linear combination of variables is used as a splitting rule.

4. CART is not at all affected by the outliers, collinearities, heteroskedasticity, or distributional error structures that affect parametric procedures. Outliers are isolated into a node and thus have no effect on splitting. Contrary to situations in parametric modeling, CART makes use of collinear variables in "surrogate" splits.

5. CART has the ability to detect and reveal variable interactions in the dataset.

6. CART does not vary under a monotone transformation of independent variables; that is, the transformation of explanatory variables to logarithms or squares or square roots has no effect on the tree produced.

7. In the absence of a theory that could guide a researcher, in a famine vulnerability study, for example, CART can be viewed as an exploratory, analytical tool. The results can reveal many important clues about the underlying structure of famine vulnerability.

8. CART's major advantage is that it deals effectively with large datasets and the issues of higher dimensionality; that is, it can produce useful results from a large number of variables submitted for analysis by using only a few import ant variables.

9. The inverted-tree-structure results generated from CART analysis are easy for anyone to understand in any discipline.

However, CART analysis does have some limitations. CART is a blunt instrument compared to many other statistical and analytical techniques. At each stage, the subdivision of data into two groups is based on only one value of only one of the potential explanatory variables. If a statistical model that appears to fit the data exists, and if its' basic assumptions appear to be satisfied, that model would be preferable, in general, to a CART tree.

A weakness of the CART method and, hence, of the conclusions it may yield is that it is not based on a probabilistic model. There is no probability level or confidence interval associated with predictions derived from a CART tree that could help classify a new set of data. The confidence that an analyst can have in the accuracy of the results produced by a given CART tree is based purely on that tree's historical accuracy — how well it has predicted the desired response in other, similar circumstances.

SVM model. Support vector machine (SVM) method which is widely used in the research area of artificial intelligence, is based on the theory of statistical learning which is concluded by combining VC dimension theory with principle of risk minimum. It is based on complexity and learning ability of the model to seek the best compromise, and finally get the best generalization ability.

Classification algorithm of SVM. SVM, which is under the condition of linear separable case, is developed from the optimal separating hyperplane. In Figure 3, we show the basic principle of SVM. On the diagram, solid and hollow points on behalf of the two kinds of samples, *H* is the classification hyperplane, *H*1 is sample that is most close to *H* and *H*2 is hyperplanes which is parallel to *H*. They have equal distance to *H*, the distance between them is termed classification interval. The hyperplane is the one that separates the two types with biggest distance, for such an ill-posed problem of classifying two kinds of samples, the optimal separating hyperplane has maximum stability and high generalization ability.

Steps involved in the design of SVM are as follows:

1. Hyperplane acting as the decision surface is defined as

$$\sum_{i=1}^{N} \alpha_i d_i K\left(x, x_i\right) = 0$$

where $K(x, x_i) = \phi^T(x)\phi(x_i)$ represents the inner product of two vectors induced in the feature space by the input vector x and input pattern xi pertaining to the example. This term is referred to as inner-product kernel.

Where

$$W = \sum_{i=1}^{N} a_i d_i \varphi\left(x_i\right)$$

$\varphi(x) = [\varphi_0(x), \varphi_1(x), \ldots, \varphi_{m1}(x)]^T$

$\varphi_0(x) = 1$ for all x

w_0 denotes the bias b.

2. The requirement of the kernel $K(x, x_i)$ is to satisfy Mercer's theorem. The kernel function is selected as a polynomial learning machine

$K(x, x_i) = (1 + x^T x_i)^2$.

3. The Lagrange multipliers $\{\alpha_i\}$ for $i = 1$ to N that maximize the objective function $Q(\alpha)$, denoted by α_0, i is determined

$$Q\left(\alpha\right) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j K\left(x, x_i\right).$$

Subject to the following constraints:

$$\sum_{i=1}^{N} \alpha_i d_i = 0$$

$$0 \leqslant \alpha_i \leqslant C \text{ for } i = 1, 2, \ldots, N.$$

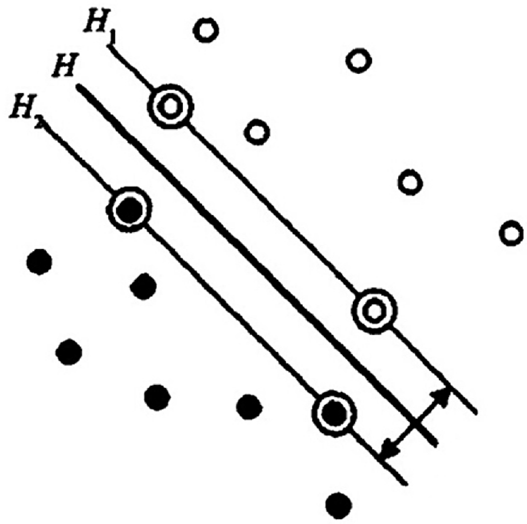


Figure 3. Optimal hyperplane for linearly separable patterns with cathegory interval = $2/\|w\|$.

4. The linear weight vector w_0 corresponding to the optimum values of the Lagrange multipliers was determined using the following formula:

$$w_0 = \sum_{i=1}^{N} \alpha_{0,i} d_i \varphi\left(x_i\right)$$

$\varphi(x_i)$ is the image induced in the feature space due to x_i; w_0 represents the optimum bias b_0.

Application of the Building Models Process to the Countries' Conditions

Data source. The statistical data of the economic and financial variables considered in this paper come from the following sources:

International Monetary Fund (World Economic Outlook database) — http://www.imf.org/external/index.htm

The World Bank (World Development Indicators database) — http://www.worldbank.org/

Exogeneous variables. I used as an exogeneous variables 387 macroeconomic factors with the coverage period from 1980 up to 2016 for 217 countries, basing on which the model has been reproduced. Concerning the open source data, should be taken into account the various density of data filling while transitioning from one factor to another. Furthermore, fullfilling of the most factors was very low, which makes the issue of their usage problematic. There are completely filled factors for the entire time period, but some of the data is partially filled.

Table 1
Lists the different country risk levels or labels used by rating agencies

Conversion of response into a numerical scale				
TE	S&P	Moody's	Fitch	
100	AAA	Aaa	AAA	Prime
95	AA+	Aa1	AA+	High grade
90	AA	Aa2	AA	
85	AA-	Aa3	AA-	
80	A+	A1	A+	Upper medium grade
75	A	A2	A	
70	A-	A3	A-	
65	BBB+	Baa1	BBB+	Lower medium grade
60	BBB	Baa2	BBB	
55	BBB-	Baa3	BBB-	
50	BB+	Ba1	BB+	Non-investment grade
45	BB	Ba2	BB	speculative
40	BB-	Ba3	BB-	
35	B+	B 1	B+	Highly speculative
30	B	B 2	B	
25	B-	B 3	B-	
20	CCC+	Caa1	CCC	Substantial risks
15	CCC	Caa2		Extremely speculative
10	CCC-	Caa3		In default with little
	CC	Ca		prospect for recovery
5	C	C		
0	D	/	DDD	In default
		/	DD	
			D	

Endogeneous variables. As a response, I used the credit ratings of countries from the leading rating agencies S&P, Moody's and Fitch. We have received a history of ratings since 1949. It is worth mentioning that in the initial period of time the process of assigning ratings, covered a limited number of countries. Only since 1990, the sample of ratings has become sufficiently representative to construct statistical classification models.

Conversion of response into a numerical scale. To quantify the data of the three rating agencies, to classify the classes, and to obtain the possibility of applying the models of classification with the models of generalized regression, the response was digitized.

We have converted the Standard & Poor rating scale (ranging from AAA to SD) into a numerical scale (ranging from 0 to 100) and shall liberally refer to both of them as S&P ratings.

Table 1 lists the different country risk levels or labels used by Standard & Poor, Moody's and Fitch and provides also descriptions associated with these labels. Countries which are assigned a label inferior to BB+ are considered as non-investment grade (speculative) countries. Countries rated CCC+ or lower are regarded as presenting serious default risks. BB indicates the least degree of speculation and CC the highest. Ratings labeled from AA to CCC can be modified by the addition of a plus or minus sign to show relative standing within the major rating categories. We consider such subcategories as separate ratings in our analysis.

Data conversion. Using the Matlab software environment, we converted the data into the following form: Country, Year, Rating, and a set of all factors.

Missing data. The vital criterion is the availability of complete and reliable statistics. We eager to avoid difficulties related to missing ratings data that could reduce the statistical significance and the scope of our analysis. Further, we exclude all the rows from the training sample, in which the values of the ratings are not specified. Eventually, the volume of the training sample was reduced to 2395 points. A vast number of variables have a small coverage, hence we exclude those factors, in which the percentage of a missing data exceeds 40. Table 7 in the Appendix lists the factors' coverage.

Correlation analysis. We conducted a correlation analysis between the 164 factors and the response, in order to select the factors that will be used in further analysis in models' constuctioning. This correlation coefficient was calculated taking into account the presence of gaps in these factors. For this purpose, a special program MATLAB was used. Next, we sorted the factors, in descending order, by reducing the absolute value of the correlation of factors.

Factors' selection for the model constructioning. The criterion is the significance of variables for estimating a country's creditworthiness. We have performed an extensive literature review which played an important role in defining the set of candidate variables for inclusion in our model. Based on the correlation matrix, economic interpretation of factors and our representations, we selected a certain number of factors for models' constructioning. It worth mentioning that within the process of factors' selection, we have omited those factors that had a close economic sense and, correspondingly, were too correlated with each other. Thus, we tried to avoid the problem associated with multicollinearity. The clarification of the factors, participated in the model constructioning are listed below. Furthermore, the correlation between factors and response has been clarified. Table 2 represents the correlation matrix between the factors and response which have been selected.

The interpretation of the correlation matrix coefficients. For the purpose of analyzing the correlation matrix, we introduce the notion of a coupling force. It is generally accepted that the strength of the correlation coefficient, as one indicator of a boundary measure, is differentiated into three levels for both positive and negative correlations. First and foremost, we will commence with analyzing the correlation coefficiencts between predictiors and response. As we can see from the correlation matrix, generally there are positive correlation coefficients as well as negative ones. According to the notion of the coupling force, we will determine the following factors as ones with the positive strong coupling force: Adjusted net national income per capita (constant 2010 US$), Household final consumption expenditure per capita (constant 2010 US$), GDP per capita (constant 2010 US$). To make the analysis clearer and concise, we will clarify the correlation of each factor in the way that is more detailed.

If to consider the further factor, Adjusted net national income per capita (constant 2010 US$), with the most strong correlation of 0.826, we may conclude, that with the slight increase of Adjusted net national income per capita, the S&P rating increases considerably. To make it easy for understanding, we will clarify that with the increase of net national income per capita, the consumption is increasing, the savings are also increasing, and therefore the economy of the country is boosted. Hence, the country becomes more reliable to invest, as the probability of debt repayment i.e. meeting the obligations increases. Therefore, the rating agencies assign more enhanced rating to this country.

Moving on to the Household final consumption expenditure per capita (constant 2010 US$), with the correlation of 0.802, we may see the similar tendency. With the growth of this indicaor, the assigned rating enhances in the same way. To make it clear, we will build the following bounderies. The Household final consumption expenditure increases, which stimulate the production, therefore the demand for goods and services is growing and the whole economy is flourishing. Hence, the country becomes more reliable to invest, as the probability of debt repayment i.e. meeting the obligations increases, which ultimately leads for country rating boost.

If to cover the last indicator in that group, namely, GDP per capita (constant 2010 US$), it is obvious, that its correlation with the response is relatively less in comparison with the above mention pair of factors, but still strong — 0.781. To make it more clear, with an increase of GDP per capita, both the consumption, the demand, the savings, the living standards are expanding. Hence,

Table 2
Correlation matrix

Factors	S&P rating	Adjusted net national income per capita (constant 2010 US$)	Household final consumption expenditure per capita (constant 2010 US$)	GDP per capita (constant 2010 US$)	Household final consumption expenditure, etc. (% of GDP)	Adjusted savings: consumption of fixed capital (% of GNI)	Urban population (% of total)	Lending interest rate (%)	Exports of goods and services (constant 2010 US$)	Broad money to total reserves ratio	General government final consumption expenditure (% of GDP)
Adjusted net national income per capita (constant 2010 US$)	0.826	1.000	0.971	0.983	-0.520	0.514	0.558	-0.420	0.466	0.321	0.436
Household final consumption expenditure per capita (constant 2010 US$)	0.802		1.000	0.938	-0.386	0.528	0.546	-0.387	0.490	0.375	0.426
GDP per capita (constant 2010 US$)	0.781			1.000	-0.541	0.470	0.559	-0.387	0.396	0.287	0.359
Household final consumption expenditure, etc. (% of GDP)	-0.563				1.000	-0.332	-0.381	0.257	-0.172	-0.034	-0.239
Adjusted savings: consumption of fixed capital (% of GNI)	0.515					1.000	0.368	-0.196	0.266	0.187	0.417
Urban population (% of total)	0.470						1.000	-0.153	0.251	0.176	0.399
Lending interest rate (%)'	-0.450							1.000	-0.280	-0.202	-0.094
Exports of goods and services (constant 2010 US$)	0.446								1.000	0.447	0.144
Broad money to total reserves ratio	0.380									1.000	0.090
General government final consumption expenditure (% of GDP)	0.379										1.000

the country becomes more reliable to invest, as the probability of debt repayment i.e. meeting the obligations increases. Therefore, the rating agencies assign more enhanced rating to this country.

Moreover, among selected facrors, we may refer the ones to the group with the positive moderate coupling force:

Adjusted savings: consumption of fixed capital (% of GNI), Urban population (% of total), Exports of goods and services (constant 2010 US$), Broad money to total reserves ratio, General government final consumption expenditure (% of GDP).

The first one, exactly, Adjusted savings: consumption of fixed capital (% of GNI), has the strongest correlation between the resonse within this group of 0.515. Consequently, we may resume, that with the increase of this indicator, the S&P rating as well increases, but in more stepless way. In order to clarify, we will build up the whole chain. When the level of savings increases, hence the consumption (demand) may decrease, as well as the production and GDP. Ultimately, the country rating is less likely to increase.

The next one, in particular, Urban population (% of total), has slightly less correlation of 0.470 in compatison with the previous one. Sequently, we may affirm that with the increase of this indicator, the rating will slightly increase as well. To make it more clear, it should be emphasized that the more urban population, the less rural one. If accept this phenomena, then we may see the logic that the overall country economy will be transferred from the agricultural to the industial type. In other words, industrialization may occur, which in turn, is going to increase the production capacity, level of consumption, living standards, volume of savings and ultimately GDP. Hence, the country becomes more reliable to invest, as the probability of debt repayment i.e. meeting the obligations increases. Therefore, the rating agencies assign more enhanced rating to this country.

The following factor, Exports of goods and services (constant 2010 US$), has the correlation of 0.446, which relatively the same as the previous one. If to clarify, with the expand of Exports of goods and services in the country, the rating will excersize minor change. To make it more clearly, we will clarify the overall chain. With the increase of Exports of goods and services, the production capacity increases, hence, the unemployment is decreasing, the living standards are enhancing and lastly the GDP is increasing. Hence, the country becomes more reliable to invest, as the probability of debt repayment i.e. meeting the obligations increases. Therefore, the rating agencies assign more enhanced rating to this country.

The last two factors from this group have approximetly the same correlation coefficients. The first one, Broad money to total reserves ratio, has relatively small correlation of 0.380. It means that with the small increase of this coefficient, the rating is less likely to be altered. To make it more clear and concise, we may clarify, that with the increase of this ratio the volume of the broad money in expanding, bringing about the inflation growth. Therefore, in response latter, the prices are going to be increased. Hence, the demand will go inverse, followed by the shrinkage of supply. Ultametly, the overall economy development is going to be imprompted, causing the ratings to be less likely enhanced.

The one and the last from this group, namely, General government final consumption expenditure (% of GDP), has the correlation of 0.379. Consequently, we may resume, that with the increase of this indicator, the S&P rating as well increases, but in more steeples way. In order to clarify, we will build up the whole chain of clarification. Due to an increase of this indicator, we may state that the general demand will undergo the growthing phase. Thus, it will stipulate the increase of supply, which in turn will lead to the expansion of the sector's of economy, ending up the economy development. Hence, the country becomes more reliable to invest, as the probability of debt repayment i.e. meeting the obligations increases. Therefore, the rating agencies assign more enhanced rating to this country.

The ultimate group of factors, which ought to be mentioned has the negative moderate coupling force. Referring to the first one, Household final consumption expenditure, etc. (% of GDP), the correlation of which is standed for –0.563. To be more consize, it means that with the increase of this indicator, the country rating will be diminished.

To make it clearly for understanding, if the consumption of the housholds is going to be increased, hence the less funds will be left for savings (putting money under deposits). Therefore, due to the shortage of the deposits, the banks will not have sufficient funds to make credit operations. Therefore, they will fall back on increasing interest rates (in the case of the regection of credit from the Central Bank), which will make the acsess for the funds more sophisticated. Ultimately, the volume of investments will be decreased harshly, slowing down the growth of the GDP. Thefore, the rating is less likely to be altered.

And, the ultimate factor, Lending interest rate (%), which has the correlation of –0.450, means that with the increase of the Lending interest rate, the country rating is to be deteriorated. In other words, the increase of the key interest rate is accompanied with the tight monetary policy which is proclaimed by the Central Bank of country. Therefore, the level of cash distribution throughout the economic agents and the economy in the whole is significantly falling down. In turn, this phenopena gives a rise to shrinkage of the volume of investment as within the country as well as from outside it.

It is worth mentioning, that according to the yield curve, increased yield is giving rise to the risk exposure, which is similar to our context with the interest rates. Ultimatly, the economy development is deterorating, country risks are going to rocket and in that case the probability of increased country rating is going down.

As we have already described the correlation between predictor and response, its time to analyze the correlation between factors as well.

We will give a rise to such notion as multicollinearity. It means that there is any mutual strong interconnection between facrots within the sample. It is worth mentioning that if it exists, therefore it deteriorates the sample with the additional boundaries between factors. Ideally, the correlation between facrors should be tended to vanish (seek a null position). In other words, they ought to be statistically independent.

If you analyze our sample, it should be clearly seen, that there is a high multicollinearity between the following factors:

1. Adjusted net national income per capita (constant 2010 US$) and Household final consumption expenditure per capita (constant 2010 US$).

2. Adjusted net national income per capita (constant 2010 US$) and GDP per capita (constant 2010 US$).

3. Household final consumption expenditure per capita (constant 2010 US$) and GDP per capita (constant 2010 US$).

The multicollinearity between these 3 pairs of facrors is almost 1, which considerable deteriorates our sample. Conducting the correlation alalysis of the data sample, we may conclude that the majority of factors have a high correlation with the response value. The highest value is presented by *Adjusted net national income per capita (constant 2010 US$)* and the least by *General government final consumption expenditure* (% of GDP) respectively.

Moreover, it can be clearly seen, that there is a high Multicollinearity (almost 1) between the following factors:

1. Adjusted net national income per capita (constant 2010 US$) and Household final consumption expenditure per capita (constant 2010 US$).

2. Adjusted net national income per capita (constant 2010 US$) and GDP per capita (constant 2010 US$).

3. Household final consumption expenditure per capita (constant 2010 US$) and GDP per capita (constant 2010 US$).

Summing up the carried out dicriptive statistics, namely, Group Means and Confidence Intervals, Discriptive Statistics for the group of factors, and Correlation matrix, we can colclude that:

On the one hand, there is the variety of advantages over the data sample, which are the following:

1. First and foremost, the majority of factors have reletevely high correlation coefficiencts with the response value.

2. Moreover, across almost all factors the intence multicollinearity is not observed.

The analysis of the relationship between factors and rating based on mean group values and confidence intervals. Eventually, with the help of conducting the descriptive statistics concerning Group Means and Confidence Intervals, we may conclude that, almost all factors have a positive correlation with the response except Household final consumption expenditure, etc. (% of GDP) and Lending interest rate (%).

In addition, due to the corresponding of one value of the factor to various confidence intervals, countries can be assigned a different rating. Thus, we may conclude that for the factor levels corresponding to different confidence intervals, the value of the rating is adjusted by other factors.

It is obvious that low rating values are prescribed by a narrow confidence interval, due to a small spread of factor values between the minimum and maximum values. Meanwhile, as the

Table 3
Discriptive statistics for the group of factors

Factors	Mean	Min	Max	STD	Median	Mode	Skew-ness	Kur-tosis
Adjusted net national income per capita (constant 2010 US$)	14377.77	277.50	66462.98	15203.35	7048.62	277.50	1.24	3.60
Household final consumption expenditure per capita (constant 2010 US$)	9496.61	288.74	41428.17	9469.80	5461.40	288.74	1.29	3.94
GDP per capita (constant 2010 US$)	18633.51	335.86	91593.67	20066.35	9002.54	335.86	1.36	4.11
Household final consumption expenditure, etc. (% of GDP)	59.35	13.41	89.27	12.62	58.86	13.41	-0.44	4.12
Adjusted savings: consumption of fixed capital (% of GNI)	13.31	2.07	24.13	4.63	13.63	2.07	-0.11	2.38
Urban population (% of total)	66.34	8.55	100.00	21.03	72.63	100.00	-0.78	2.95
Lending interest rate (%)	12.55	0.50	86.36	9.35	10.48	5.38	3.15	17.82
Exports of goods and services (constant 2010 US$)	123303072753	494277379	2208569650600	192036214575	59391883047	494277379	4.72	40.77
Broad money to total reserves ratio	7.61	0.24	90.94	12.53	3.49	0.24	3.66	17.32
General government final consumption expenditure (% of GDP)	15.99	5.69	27.63	4.75	16.00	5.69	0.17	2.19

rating increases, the spread of factor values rises as well, which leads to wide confidence intervals.

By no means unimportant is that due to the peculiarity of the data sample, for a high rating values, namely from 80 to 90, there is a low number of observations for this factor, which leads to fluctuations for both the mean value and confidence interval.

The summary of carrying out the discriptive statistics for the group of factors. Conducting the descriptive analysis for the group of factors, we may conclude that almost all factors has multimodal distribution (has two modes). Consequently, we may conclude that it is far from the normal distribution.

Besides, the sample has right-sided asymmetry, as evidenced by a positive Skewness. In other words, the mean value and the median one is greater than the mode value. This suggests that our data is shifted to the left, relative to the normal distribution. Consequently, there is an abnormality in the distribution of the values of factors.

It should not go unspoken about the fact, there is an excess in the data sample. Fot the major part of data sample, the peak values are higher than the normal distribution. This is confirmed by a value of Kurtosis more then 3 whereas the normal value is 3.

However, the following factors are differ from the data sample: Household final consumption expenditure, etc. (% of GDP), Adjusted savings: consumption of fixed capital (% of GNI), Urban population (% of total).

The distribution is unimodal for this group of factors (has one mode). Consequently, we may conclude that it is close to the normal distribution.

However, the sample has left-sided asymmetry, as evidenced by a negative value of Skewness. In other words, in comparison with previous factors, the mean value and the median one is not much but still greater than the mode value. This suggests that our data is slightly shifted to the right, relative to the normal distribution. Consequently, there is a normality in the distribution of the values of factors.

Furthermore, there is an excess in the data sample. The peak values are less then normal distribution. This is confirmed by a value of kurtosis less then 3, whereas the normal value is 3.

Conducting the tests on normality. Based on our assumption of non-parametricity of data, we are going to test our data on normality, using the number of tests. The normality assumption is at the core of a majority of standard statistical procedures, and it is important to be able to test this assumption. In addition, showing that a sample does not come from a normally distributed population is sometimes of importance *per se*. Among the procedures used to test this assumption, one of the most well known is a Kolomogorov–Smirnov test.

First, we verify the null hypothesis concerning the normality of distribution using The Kolmogorov–Smirnov test. In Matlab application, we used the *ks*-test function. It returns a test decision for the null hypothesis that the data in vector *x* comes from a standard normal distribution, against the alternative that it does not come from such a distribution, using the one-sample Kolmogorov–Smirnov test. The result *h* is 1 if the test rejects the null hypothesis at the 5% significance level, or 0 otherwise.

The findings of the Kolmogorov–Smirnov test showed that throughtout the sample, all factors reject null hypothesis. Therefore, we may conclude that there is an absence of normality and our sample has non-normal distribuion.

Moreover, we conducted the additional test to verify the null hypothesis concerning the normality of distribution using the Anderson–Darling test. In Matlab application, we used the ad-test function. It returns a test decision for the null hypothesis that the data in vector *x* is from a population with a normal distribution, using the Anderson–Darling test. The alternative hypothesis is that *x* is not from a population with a normal distribution. The result *h* is 1 if the test rejects the null hypothesis at the 5% significance level, or 0 otherwise.

The findings of the Anderson-Darling test showed the same results as Kolmogorov–Smirnov. To be more clear and concise, it confirmed that throughtout the sample, all factors reject null hypothesis. Therefore, we may conclude that there is an absence of normality and our sample has non-normal distribuion.

Construction and verification of models of classification. In our scientific paper, we have been using both the special application, namely Matlab Toolbox, and writining the code by ourself. Let us consider the Matlab Toolbox in more detail.

Construction of models, including missing data. After selecting and preparing a database of factors and responses, the first model was created using the "Classification Learner". The latter allows solving classification problems and building models in an interactive mode.

Using this approach, the training sample size is 2395 points.

Method 1. Classification and regression trees. The methods of classification and regression trees showed the following results:

The accuracy of various types of classification and regression trees is shown in the table 4.

Based on the accuracy analysis of the methods, we can conclude that Bagged Trees showed the highest accuracy of 51.5%. Therefore, we take this type of classification and regression trees as a basis in subsequent analyzes and model constructions.

After matrix analyzing, we can conclude that this model is prone to overvaluation of the rating, if to be concrete, it overestimates the set of points and assigns them a ration equal to 100 or AAA. In addition, we see that for the critical cases, namely the countries of bankruptcy or close to them, the model does not determine it (0, 10, 15, 20).

Method 2. Discriminant Analysis. The methods of discriminant alalysis showed the following results:

Proceeding from the table 5, we will draw the following conclusions:

1. Linear discriminant analysis showed a classification accuracy of 23.6%, which indicates about significant classification problems.
2. Unfortunately, quadratic discriminant analysis failded due to degeneracy of the covariance matrix which is caused by the strong collinearity between factors.
3. Subspace discriminant showed the best accuracy among other types of discriminant analysis at 28.5%, but this accuracy is too small for classification.

After matrix analyzing, we can conclude that this model is prone to overvaluation of the rating, if to be concrete, it overestimates the set of points and assigns them a ration equal to 100 or AAA. In addition, we see that for the critical cases, namely, the bankrupt countries or close to them, the model does not determine it (0, 10, 15, 20). A very strong deviation from the true value is also noticeable, which indicates a poor quality of classification. The general conclusion: discriminant analysis gives an unsatisfactory result.

Method 3. Support Vector Machine (SVM). The methods of SVM showed the following results:

Proceeding from the table 6, we will draw the following conclusions:

1. Based on the accuracy analysis of the methods, we can conclude that Cubic SVM showed the highest accuracy of 28.5%. Therefore, we take this type of SVM as a basis in subsequent analyzes and model constructions.
2. After matrix analyzing, we can conclude that this model is prone to undervaluation of the rating, if to be concrete, it underestimates the set of points and assigns them a ration equal to 0 (D).

Table 4
The accuracy of various types of classification and regression trees

Type of Classification and Regression Tree	Accuracy (%)
Simple Tree	24.6
Medium Tree	32.1
Complex Tree	38.9
Boosted Trees	33.4
Bagged Trees	51.5
RUSBoosted Trees	26.0

Table 5
The accuracy of various types of discriminant analysis

Type of Discriminant Analysis	Accuracy (%)
Linear Discriminant	24.8
Quadratic Discriminant	Failed
Subspase Discriminant	28.5

Table 6
The accuracy of various types of SVM

Type of SVM	Accuracy (%)
Linear SVM	20.7
Quadratic SVM	26.6
Cubic SVM	28.5
Fine Gaussian SVM	28.1
Medium Gaussian SVM	24.4
Coarse Gaussian SVM	16.8

Table 7
The accuracy of various types of k-NN

Type of k-NN	Accuracy (%)
Fine k-NN	41.7
Medium k-NN	33.9
Coarse k-NN	25.6
Cosine k-NN	33.4
Cubic k-NN	33.6
Weighted k-NN	40.9
Subspace k-NN	24.8

3. A very strong deviation from the mean value is also noticeable, which indicates a poor quality of classification.

The general conclusion: SVM gives an unsatisfactory result.

Method 4. k-nearest neighbors (k-NN) The methods of *k*-NN showed the following results.

Proceeding from the table 7, we will draw the following conclusions:

1. Based on the accuracy analysis of the methods, we can conclude that Fine *k*-NN showed the highest accuracy of 41.7%. Therefore, we take this type of *k*-NN as a basis in subsequent analyzes and model constructions.

2. After matrix analyzing, we can conclude that this model is prone to overvaluation of the rating, if to be concrete, it overestimates the set of points and assigns them a ration equal to 100 or AAA. In addition, we see that for the critical cases, namely the countries of bankruptcy or close to them, the model does not determine it (0, 10, 15, 20).

3. A very strong deviation from the true value is also noticeable, which indicates a poor quality of classification.

The general conclusion: the model sufficiently deviates from the true values, but the accuracy estimates are at least better than in the Discriminant Analysis and the Support Vector Mashine Method. As a result, it can be concluded that the majority of builded models with presence of missing data (NAN) in the original sample show insufficient accuracy; hence, this methodology cannot be applied.

However, it can be clearly seen that the method of classification and regression trees, as well as the method of the nearest neighbor, showed quite good results. Consequently, using the following approach, we will choose only the following methods, specifically, ones with the highest accuracy:

1. Classification and Regression Tree: Bagged Trees.
2. Discriminant Analysis: Subspase Discriminant.
3. Support Vector Machine: Cubic SVM.
4. *k*-nearest neighbors: Fine *k*-NN.

Building Models with Exclusion of the Missing Data

When using this methodology, we exclude all points in which missing data exists at least in one factor. As a result, our training sample shrinked from 2395 to 1042 points. On the one hand, we have got rid of missing data, which positively influences the construction of the models, but on the other, we lost a certain amount of information.

Method 1. Classification and regression trees: Bagged Trees. The results of building the model were shown as a Confusion Matrix Graph. If to compare with the previous approach, we can see a considerable improvement, particularly, the accuracy of the classification has increased from 51.5% to 70.5%. Also when analyzing the matrix, we see a significant enhancement in the accuracy of the classification. The points deviate much less from the true value. However, the model has a significant drawback. It was impossible to classify properly critical countries with low ratings (bankrupt).

Method 2. Discriminant Analysis: Subspase Discriminant. The results of building the model were shown as a Confusion Matrix Graph. If to compare with the previous approach, we see that the accuracy of the model has improved, but not considerably. The accuracy of classification has increased from 28.5% to 31.5%. When analyzing the matrix, we did not notice significant improvements in comparison with previously built models. Also as with classification and regression trees, the model has a significant drawback. It was impossible to classify properly critical countries with low ratings (bankrupt).

Method 3. Support Vector Mashine: Cubic SVM. The results of building the model were shown as a Confusion Matrix Graph. If we compare it with the previous approach, we can see a considerable improvement, particularly, the accuracy of the classification has increased from 28.5% to 68.2%. Also, when analyzing the matrix, we see a significant enhancement in the accuracy of the classification. The points deviate much less from the true value. However, the model has a significant drawback. It was impossible to classify properly critical countries with low ratings (bankrupt).

Method 4. *k*-nearest neighbors: Fine *k*-NN. The results of building the model were shown as a Confusion Matrix Graph. If to compare with the previous approach, we can see a considerable improvement, particularly, the accuracy of the classification has increased from 41.7% до 70.2%. Also, when analyzing the matrix, we see a significant enhancement in the accuracy of the classification.

Table 8
The accuracy of the models with and without missing data

Type of method	Accuracy of the models including missing data (%)	Accuracy of the models with exclusion of the missing data (%)
Bagged Trees	51.5	70.5
Subspace Discriminant	28.5	31.5
Cubic SVM	28.5	68.2
Fine KNN	41.7	70.2

The points deviate much less from the true value. However, the model has a significant drawback. It was impossible to classify properly critical countries with low ratings (bankrupt).

After conducting this approach, it can be concluded that all models have enhanced their results. The greatest increase in accuracy was obtained by the Bagged Trees, Cubic SVM and Fine KNN. Meanwhile Subspace Discriminant has not shown significant improvements.

Analysis of uniformity of distribution of sampling points by classes. It can be concluded that almost all models have enhanced their accuracy, but unfortunatuly, all of them unable to classify properly critical countries, ones with low ratings (bankrupt), which are vital for us.

We are convinced that this was due to the uneven presence of countries with various ratings within the training sample, specifically: the moments when countries are assigned low ratings are relatively rare, meanwhile for countries with the highest rating, almost all points are generally known for all factors. Consequently, points with high ratings dominate in our training sample.

Below is the histogram of the distribution of our sample by classes.

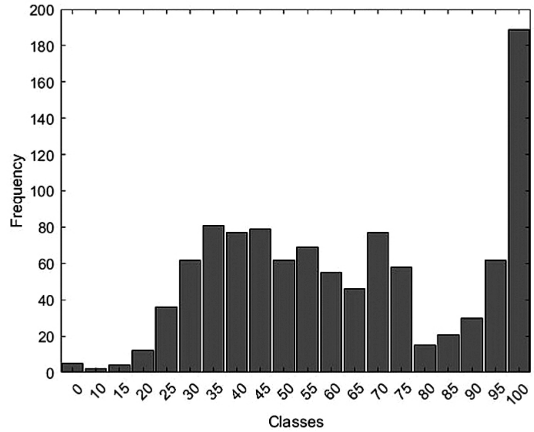


Figure 4. The histogram of the distribution of our sample by classes.

The histogram shows that in classes that correspond to countries with a low rating (from 0 to 25), there is a very small number of points, whereas in the class that corresponds to the highest rating (100), there is a prevailing number of points. The highlighted fact strongly distorts the construction of the model.

The method of grouping and converting data in manual mode. Applying this method, we eliminated 30 points from the group with the highest rating. Moreover, we also combined groups with a low rating in the range (from 0 to 25) to one group, and set its value to 10.

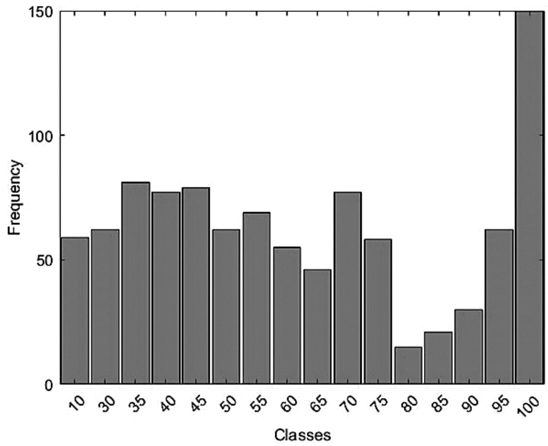


Figure 5. Classes distribution after aggregation.

The histogram clearly shows that the frequency of low-rated countries has increased significantly. Also, the quantity of points in the class with the highest rating was reduced up to 150.

Method 1. Classification and Regression Trees: Bagged Trees. The use of grouping and data conversion approach did not effect significantly on the accuracy improvement in the model. The accuracy of the model increased from 70.5% to 72%. However, the graph shows that the amalga-

Table 9
The accuracy of the models without missing data and with grouping and data conversion approach

Type of method	Accuracy of the models with exclusion of the missing data (%)	Accuracy of the models with grouping and data conversion approach (%)
Bagged Trees	70.5	72.0
Subspace Discriminant/Quadratic Discriminant	31.5	51.2
Cubic SVM	68.2	66.1
Fine KNN	70.2	72.4

mation of groups with a low rating (from 0 to 25), and a reduction in sampling points for countries with high ratings, had a positive effect on the model, specifically decreasing deviations from the true value. The model commenced assigning values to the class of countries with poor rating. In addition, the model has become less inclined to overestimate the values of the groups of countries with the highest rating.

Method 2. Discriminant Analysis: Quadratic Discriminant. The results of building the model are shown below as a Confusion Matrix Graph. Using this approach, the method of discriminant analysis has been changed from Subspace Discriminant on Quadratic Discriminant. This has led to significant improvement in the accuracy of the model, namely from 31.5% to 51.2%. However, the graph shows that the amalgamation of groups with a low rating (from 0 to 25), and a reduction in sampling points for countries with high ratings, had a positive effect on the model, specifically decreasing deviations from the true value. The model commenced assigning values to the class of countries with poor rating. In addition, the model has become less inclined to overestimate the values of the groups of countries with the highest rating.

Method 3. Support Vector Mashine: Cubic SVM. The results of building the model were shown below as a Confusion Matrix Graph. The use of grouping and data conversion approach did not effect significantly on the accuracy improvement in the model, but on the contrary to its deterioration. The accuracy of the model decreased from 68.2% to 66.1%. However, the graph shows that the amalgamation of groups with a low rating (from 0 to 25), and a reduction in sampling points for countries with high ratings, had a positive effect on the model, specifically decreasing deviations from the true value. The model commenced assigning values to the class of countries with poor rating. In addition, the model has become less inclined to overestimate the values of the groups of countries with the highest rating.

Method 4. *k*-nearest neighbors: Fine *k*-NN. The results of building the model are shown below as a Confusion Matrix Graph. The use of grouping and data conversion approach did not effect significantly on the accuracy improvement in the model. The accuracy of the model has grown from 70.2% to 72.4%. However, the graph shows that the amalgamation of groups with a low rating (from 0 to 25), and a reduction in sampling points for countries with high ratings, had a positive effect on the model, specifically decreasing deviations from the true value. The model commenced assigning values to the class of countries with poor rating. In addition, the model has become less inclined to overestimate the values of the groups of countries with the highest rating.

Ultimetaly, it should be concluded, that the use of grouping and data conversion approach did not effect significantly on the accuracy improvement throughout all models (Table 9). Approximetely, in average, the accuracy has been increased by 3%. However, the amalgamation of groups with a low rating (from 0 to 25), and a reduction in sampling points for countries with high ratings, had a positive effect on the models, specifically decreasing deviations from the true value. Models commenced assigning values to the class of countries with poor rating. In addition, the models has become less inclined to overestimate the values of the groups of countries with the highest rating.

Logarithmic scaling of factors. Having carried out a logarithmic scaling of factors, it can be concluded that this method did not enhance the accuracy of the model.

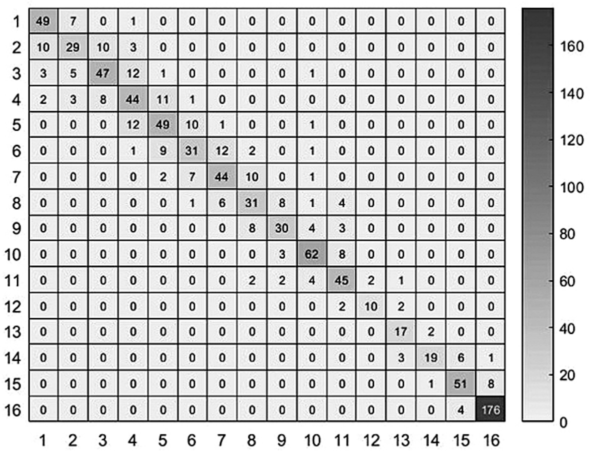


Figure 6. Confusion matrix — Ensamble tree model.

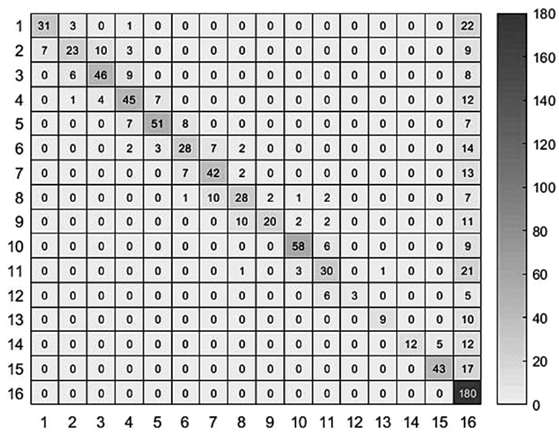


Figure 7. Confusion matrix for SVM classify model.

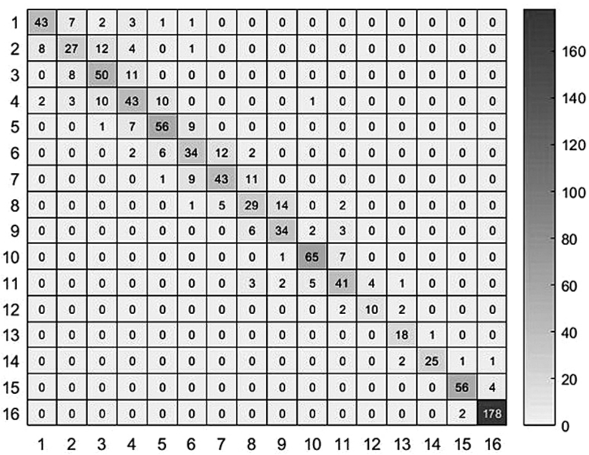


Figure 8. Confusion matrix for Fine k-NN model.

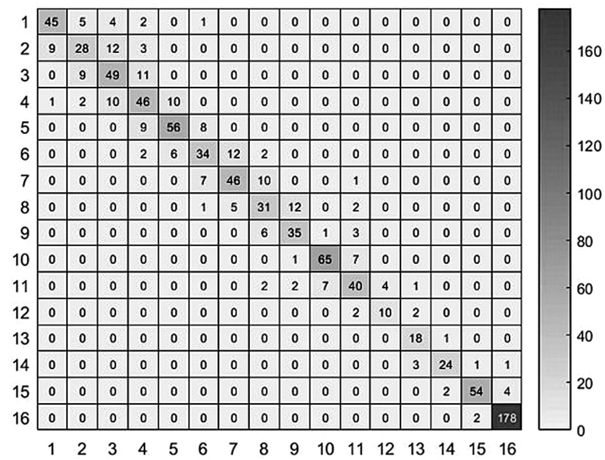


Figure 9. Confusion Matrix for enhanced k-NN model.

Dynamics of factors with the account of the previous year. So far, we have used cross-sectional models with data. In other words, the factors'data were fixed for the year of rating. Thus, we built static models. The following step in our studies was to analyze the impact of changes in factors over time on the country's rating.

For this purpose, we amalgamate the data of the current year with the data of the previous one, for the selected factors. As a result, each factor will represent a model with two vectors, the values of the current and previous year.

Moreover, we refused to use Matlab Toolbox due to some limitations and lack of functions. In our vision, Matlab Toolbox is limited in terms of determining the quality criteria of the model. In other words, Matlab Toolbox allows you to determine only the criterion of classification accuracy.

However, in the future we are going to conduct a thin comparison of the selected models. For this purpose, we will need those criteria that will evaluate classification errors, for example Mean Absolute Error (MAE) and Mean Squared Error (MSE).

As a result, a special program was created, the purpose of which was to carry out a thin comparison of models using various criteria for both the accuracy of the classification and additional criteria MAE (Mean Absolute Error) and MSE (Mean Squared Error).

Method of Classification with the Account of the Previous Year

Method 1. Classification and Regression Trees: Bagged Trees. The results of building the model were shown in figure 6 as a Confusion Matrix Graph.

This approach, namely, taking into account the previous year, significantly improved the accuracy of the model. The accuracy of classification has increased from 72% to 75.9%. Also

Table 10
Comparative data for results of the models used

Type of method	Accuracy of the models with grouping and data conversion approach (%)	Accuracy of the models taking into account the previous year (%)	Mean Absolute Error
Bagged Trees	72.0	75.9	1.8666
Quadratic Discriminant	51.2	NaN	NaN
Cubic SVM	66.1	67.1	8.6970
Fine KNN	72.4	77.7	1.7839

when analyzing the matrix, we see a significant enhancement in the deviations. The points deviate much less from the true value. The absolute error was 1.8666, which does not exceed the conventional boundary between classes, equal to 5.

Method 3. Support Vector Machine: Cubic SVM. The results of building the model are shown in figure 7 as a Confusion Matrix Graph.

This approach failed to improve significantly the accuracy of the classification, which slightly increased from 66.1% to 67.1%. Also, when analyzing the matrix, we did not notice any obvious enhancement with respect to deviations. The absolute error was 8.6970, which indicates a significant deviation from the true value for an erroneous classification.

Method 4. *k*-nearest neighbors: Fine *k*-NN. The results of building the model were shown in figure 8 as a Confusion Matrix Graph.

This approach, namely, taking into account the previous year, significantly improved the accuracy of the model. The accuracy of classification has increased from 72.4% to 77.77%. Also when analyzing the matrix, we see a significant enhancement in the deviations. The points deviate much less from the true value. The absolute error was 1.7839, which does not exceed the conventional boundary between classes, equal to 5. Using the methodology with the account of the previous year, we may conclude, that generally, the models has been enhanced significantly.

Particularly, the following models showed perfect response:

Bagged Trees

The accuracy of the classification has increased significantly, namely by 6%. And moreover, using our own algoritm, we are aware of the value of mean absolute error, which equals 1,86. Therefore, we may conclude that the deviation is too low, bacause it does not exceed the conventional boundary between classes, equal to 5.

Table 11
Results for enhanced k-NN model with two distance metrics

Metrics type	Accuracy	Mean Absolute Error
Euclidean	0.7777	1.7839
Manhattan	0.7849	1.6339

Fine *k*-NN

The next and one model, which also showed perfect results is Fine *k*-NN. The accuracy of the classification has increased significantly, approximately by 6%. Moreover, using our own algoritm, we become aware of the value of mean absolute error, which equals 1.78. Therefore, we may conclude that the deviation is too low, because it does not exceed the conventional boundary between classes, equal to 5. Unfortunalely, Quadratic Discriminant is failed this test.

Cubic SVM

Cubic SVM has also enhanced its results, but insignificantly. The accuracy of the classification has increased only by 1%, which can claim that this method does not fit to this approach. Furthermore, the value of mean absolute error is too high, particularly, 8.69, which exceeds the conventional boundary between classes, equal to 5.

Enhancing of *k*-NN methodology

In order to improve the accuracy of the classification, besides the Euclidean metrics, we verified, and other ones. As a result, the metric

CityBlock or Manhattan showed the best result, enhancing the accuracy of the model by 1%, which is a significant improvement. The results of building the model were shown in figure 9 as a Confusion Matrix Graph.

Conclusions

Carrying out the relatively vast number of tests and constructioning a significant quantity of models, we conclude that the most appropriate are the following two: Classification and Regression Trees (Bagged Trees) and *k*-nearest neighbors: Fine *k*-NN.

The quantitive characteristics of the building models are given below:

The best accuracy which has been shown by the *k*-nearest neighbors, (Fine *k*-NN), using Manhattan metrics is 0.7849. Therefore, the points almost do not deviate from the true value.

The best mean Absolute Error has been shown by the *k*-nearest neighbors. Its value is 1.6339, which does not exceed the conventional boundary between classes of 5.

The one and the last model, with relatively the same characteristics, is Classification and Regression Trees (Bagged Trees).

The accuracy shown by the Classification and Regression Trees (Bagged Trees) is 75.9%. Therefore, the points almost do not deviate from the true value.

The best mean Absolute Error has been shown by the Classification and Regression Trees (Bagged Trees) is 1.8666, which does not exceed the conventional boundary between classes of 5.

In addition, the models are not inclined to overestimate the values of the groups of countries with the highest rating.

Moreover, our models have a significant adantage. They are able to classify critical countries properly, ones with low ratings (bankrupt).

Furthermore, it is vital to enlist the characteristics of the data sample used to construct models.

There is a variety of advantages over the data sample, which are the following:

First, and foremost, the majority of factors have reletevely high correlation coefficiencts with the response value. Second, across almost all factors the intense multicollinearity was not observed.

On the other hand, there are a number of significant drawbacks, which deteriorates the whole sample:

1. Almost all factors has multimodal distribution (has two modes). Consequently, we may conclude that our sample is far from the normal distribution.

2. The vital thing is that, the sample has right-sided asymmetry, as evidenced by a positive Skewness. In other words, the mean value and the median one is greater than the mode value. This suggests that our data is shifted to the left, relative to the normal distribution. Consequently, there is an abnormality in the distribution of the values of factors.

3. Equally important is the fact that there is an excess in the data sample. Fot the major part of data sample the peak values are higher than the normal distribution. This is confirmed by a value of kurtosis more then 3.

4. Based on the interpreted results of the descriptive statistics, we should sum up that our data is far from the normal distribution. In other words, the Gauss-Markov conditions are not meeted.

To conclude, we may say, that such classification techniquies and algoritms are recommended for analysis, in the absence of the insider information and small variety of factors (sample).

Sources

http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1073&context=econ_fac

https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Discriminant_Analysis.pdf

http://download.portalgaruda.org/article.php?article=81950&val=160

http://www.ics.uci.edu/~welling/teaching/273ASpring10/AdaBoost4Stocks.pdf

http://pdf.usaid.gov/pdf_docs/PNACH725.pdf

http://www.tsijournals.com/articles/research-of-risk-assessment-and-prediction-based-on-support-vector-machine.pdf

Страновой риск в международных инвестициях: структура и методы расчета

Алексей Ивкин[1]

В данной статье исследуются вопросы оценки странового риска в разрезе экономической безопасности и устойчивости экономики. Основными объектами исследования являются страновой риск и его структурные компоненты. Основная цель научной статьи — проанализировать методы оценки странового риска с разных точек зрения и предложить модель для измерения странового риска, которая позволила бы адекватно оценить страновой риск, экономическую безопасность, уровень и динамику экономической устойчивости, включая структурные компоненты и их отношения. В статье выделено несколько основных задач:

- понять важность оценки странового риска в контексте развивающихся мировых рынков, учитывая причины и элементы странового риска на основе других научных исследований;
- изучить и выяснить преимущества и недостатки методов оценки страновых рисков, а также определить способы управления рисками;
- применить количественные и качественные методы анализа, сформулировать, создать и представить модель оценки странового риска в контексте экономической безопасности и устойчивости, которая будет определять факторы, влияющие на страновой риск, и определять их прямые и косвенные отношения между собой;
- проверить практическую пригодность модели оценки странового риска путем проведения эмпирического анализа во всем мире, определения направлений смягчения последствий странового риска.

Ключевые слова: анализ странового риска; кредитные рейтинги; долг; страктурные качественные методы; дискриминантный анализ; регрессия; классификация; метод ближайших соседей; метрики расстояния; классификационные и регрессионные деревья; опорные вектора; дискриминантный анализ; недостающие данные; тест Колмогорова–Смирнова; тест Андерсона–Дарлинга; непараметрические методы

JEL classification: C53, O16, O47

[1] Студент Финансового университета, магистерская программа «Международные финансы», Москва, Россия; vok-zal@mail.ru

Научный руководитель — А.И. Ильинский, доктор технических наук, профессор кафедры «Макроэкономическое регулирование», Финансовый университет, Москва, Россия.

Review of Business and Economics Studies
2018, Vol. 6, No. 1, 78–82

DOI: 10.26794/2308-944X-2018-6-1-78–82

Business Valuation of Nike, Incorporated

Elena Miroshina (Silantieva)[1], Egor Romanov[2]

[1]Department of International Economy and World Finance, Financial University, Moscow, Russia; EASila@yandex.ru
[2]International Finance Faculty, Financial University, Moscow, Russia; egor-romanov98@mail.ru

Abstract
In this article we focused on the ability to analyze the market and to carry out complex assessment of the company, that is one of the most paramount tasks before financiers. We tried to understand if Nike, Inc. company is overvalued or not. In this work we will consider the analysis of Nike, Inc. by means of Discrete Cash Flow (DCF) as a method and the analysis of financial statements.
Keywords: business valuation; DCF; Financial Analysis; Nike, Inc.
JEL classification: D22, F23, G32

NIKE, incorporated on September 8, 1969, is engaged in the design, development, marketing and selling of athletic footwear, apparel, equipment, accessories and services. The Company's operating segments include North America, Western Europe, Central & Eastern Europe, Greater China, Japan and Emerging Markets. The Company's portfolio brands include the NIKE Brand, Jordan Brand, Hurley and Converse. The Company sells its products to retail accounts, through its retail stores and Internet Websites, and through a mix of independent distributors and licensees across the world. The Company's products are manufactured by independent contractors.

Financial Analysis

Revenues

For the first time in a while, NIKE reported flat revenue growth in its latest quarterly report. The company is facing strong competition from Adidas and Under Armour. Both competitors are going after sportswear and women segments where fashion is an important element.

Declining sales in North America along with currency headwinds explain recent revenue deception. The company is going through a transformation where it will increase direct-to-customer sales through owned stores and online sales. The idea is to get closer to clients, identify their needs and change of taste to make modification accordingly. I think this is a smart idea. Customers are shifting their buying patterns and online sales will help gather additional data to improve the company's product. Nike also started a pilot project with Amazon. As opposed to classic retailers, Nike has nothing to fear from Amazon. It's impossible for them to sell Nike shoes without Nike.

Earnings

As previously mentioned, strong competition led Nike to cut prices and spend a fortune in marketing to keep its market share. Nike is an icon brand and its marketing budget is unmatched in the industry. The company dominates the basketball industry and is taking a serious foot inside the growing soccer market. Unfortunately, as competition grows, margins are under pressure. We can see that it is slightly weaker than usual:

It is expected that Nike will increase direct sales to improve margins in the future. As the company will learn more about its clients, it should target their needs and improve the amount of sales and profits.

Dividend Growth Perspective

Nike has increased its dividend for 15 consecutive years. This makes it part of the Dividend Achievers list. The Dividend Achievers Index refers to all public companies that have successfully increased

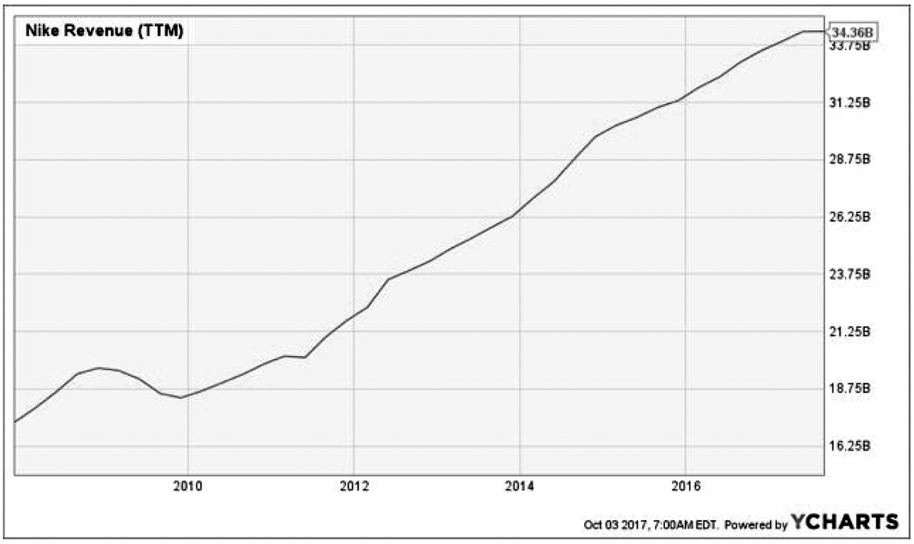


Figure 1. Nike Revenue 2010–2016.

Source: https://ycharts.com/store/nike

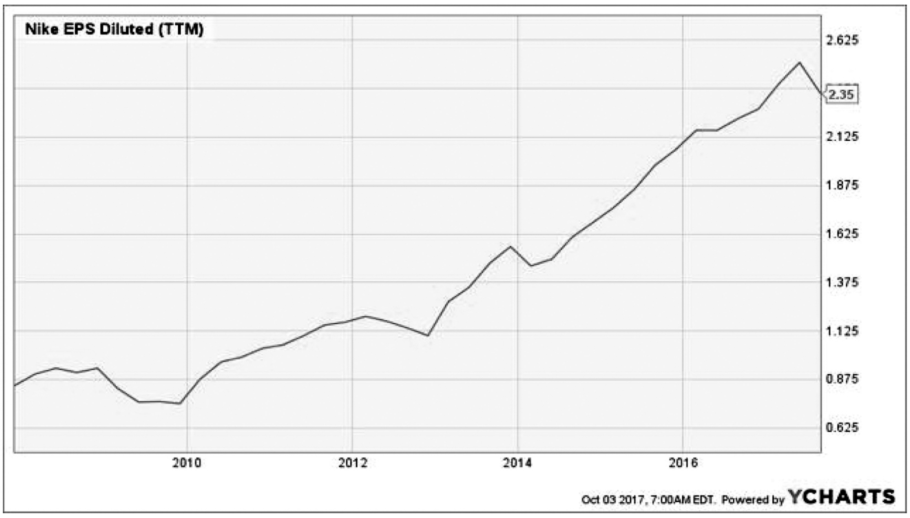


Figure 2. Nike earnings (EPS) 2010–2016.

Source: https://ycharts.com/store/nike

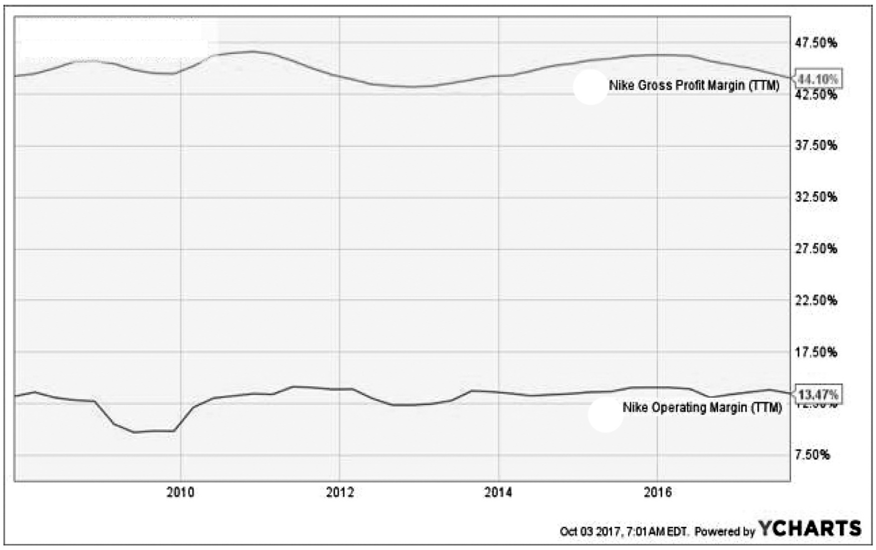


Figure 3. Nike's operating margin 2010–2016.

Source: https://ycharts.com/store/nike

Figure 4. Nike's Dividends 2008–2016.

Source: https://ycharts.com/store/nike

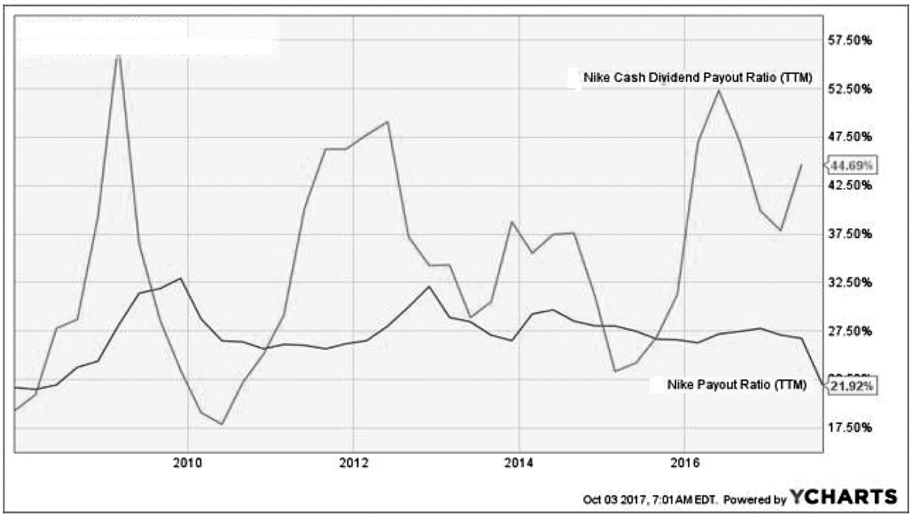


Figure 5. Nike's Payout Ratio (TTM) and Nike's Cash Dividend Payout Ratio 2010–2016.

Source: https://ycharts.com/store/nike

their dividend payments for at least 10 consecutive years. At the time of writing this work, we have found 265 companies that achieved this milestone.

The company usually flies under the radar of many dividend growth investors as its yield has not been a stock highlight for the past seven years. However, the dividend payment doubled over the past five years and the company is in line to announce its 16th consecutive dividend raise later in 2017.

The good thing about sports apparel is that the more client uses them, the more he/she buys. It's virtually impossible to keep the same running shoes for more than a year if somebody runs 3–4 times a week. This generates repetitive sales leading to continuous cash flow generation.

While management is on a roll to buy back shares and increase significantly its distribution, both payout and cash payout ratios are well under control. In fact, management has enough room to keep a high single-digit to double digit dividend growth rate for many years to come.

Potential Downsides

The company is well implemented across the world. While there are still growth vectors coming from China and emerging markets, North American and European markets may continue to go sideways. Nike has been known for its affiliation with many stars in the sport industry. Sponsoring has been a proven, but expensive strategy. As competition intensifies, such spending will hurt margins. Overall, there are not any major dark clouds hovering over Nike's head.

Discrete Cash Flow method

Let's start with the calculation of Weighted Average Cost of Capital (WACC — see https://www.gurufocus.com/term/wacc/NKE/WACC/Nike+Inc):

WACC = E / (E+D) × Cost of Equity + D / (E+D) × Cost of Debt × (1 — Tax Rate)

1. Weights:

Generally speaking, a company's assets are financed by debt and equity. We need to calculate the weight of equity and the weight of debt.

The market value of equity (E) is also called "Market Cap (M)". As of today, Nike's market capitalization (E) is $ 98217.390 million.

The market value of debt is typically difficult to calculate, therefore, we use book value of debt (D) to do the calculation. It is simplified by adding the latest two-year average Current Portion of Long-Term Debt and Long-Term Debt & Capital Lease Obligation together. As of Aug., 2017, Nike's latest two-year average Current Portion of Long-Term Debt was $ 2732 million and its latest two-year average Long-Term Debt & Capital Lease Obligation was $ 2732 million. The total Book Value of Debt (D) is $ 2920 million.

a) weight of equity = E / (E + D) = 98217.390 / (98217.390 + 2920) = 0.9711

b) weight of debt = D / (E + D) = 2920 / (98217.390} + 2920) = 0.0289

2. Cost of Equity:

We use Capital Asset Pricing Model (CAPM) to calculate the required rate of return. The formula is:

Cost of Equity = Risk-Free Rate of Return + Beta of Asset × (Expected Return of the Market — Risk-Free Rate of Return)

a) We use 10-Year Treasury Constant Maturity Rate as the risk-free rate. It is updated daily. The current risk-free rate is 2.37000000%.

b) Beta is the sensitivity of the expected excess asset returns to the expected excess market returns. Nike's beta is 0.51.

c) (Expected Return of the Market — Risk-Free Rate of Return) is also called market premium. We consider market premium to be 6%.

Cost of Equity = 2.37000000% + 0.51 × 6% = 5.43%

3. Cost of Debt:

We use last fiscal year end Interest Expense divided by the latest two-year average debt to get the simplified cost of debt.

As of May, 2017, Nike's interest expense (positive number) was 59 million. Its total Book Value of Debt (D) is $ 2920 million.

Cost of Debt = 59 / 2920 = 2.0205%.

4. Multiply by one minus Average Tax Rate:

We use the latest two-year average tax rate to do the calculation. The latest Two-year Average Tax Rate is 15.945%.

Weighted Average Cost of Capital (WACC) for today is calculated as:

WACC = E / (E+D) × Cost of Equity + D / (E + D) × Cost of Debt × (1 — Tax Rate)

WACC = 0.9707 × 5.43% + 0.0293 × 2.0205% × (1–15.945%) = 5.34%

After calculating WACC we should find out more information about the Free Cash Flow (FCF). This information is available in Bloomberg terminal.

Free Cash Flow for the future 5 year period (in millions) (Nike Inc., n.d.):

1) FCF 2017–2535

2) FCF 2018–3194

3) FCF 2019–3550

4) FCF 2020–3970

5) FCF 2021–4046.

Now we need discount all the future FCF to current day with the help of WACC.

$[FCF_{17}/(1 + WACC)] + [FCF_{18}/(1 + WACC)^2] + (FCF_{19}/(1 + WACC)^3] + (FCF_{20}/(1 + WACC)^4] + (FCF_{21}/(1 + WACC)^5] + TV/(1 + WACC)^6$ that is

2535/1.0534 + 3194/1.109 + 3550/1.168 + 3970/1.229 + 4046/1.295 + TV/1.361 = Market Capitalization.

We can calculate Terminal Value (TV) in two ways:

TV by exit multiple approach

$TV=EBITDA_{2022} \times EV/EBITDA_{2022}$ + Cash — DebtI dare to forecast that EV/EBITDA in 2022 will be 12.696 due to historical data and Bloomberg terminal.

So, TV= 8444 × 12.696 + 5519–3814= 108910 million.

Market Capitalization:

2535/1.0534 + 3194/1.109 + 3550/1.168 + 3970/1.229 + 4046/1.295 + 108910/1.361= 94765 million.

Gordon Growth Model

2.3% — Long term growth rate

$TV= (FCF_{2021} \times 1.023)/(WACC — 0.023)$

TV= (4046 × 1.023)/(0.0534–0.023)= 4139/0.034= 136151

Market Capitalization:

2535/1.0534 + 3194/1.109 + 3550/1.168 + 3970/1.229 + 4046/1.295 + 136151/1.361= 114718 million.

The financial analysis of the company has demonstrated that Nike, Inc. performs really well over the last years in the market of sport apparel and how potentially beneficial it can be in the near future. Comparative approach has shown that the market value of Nike's shares are overvalued, but it is connected with many different factors that take place in the sphere of sport apparel. However, basing on the DCF analysis the company is undervalued to its fair value, so that its shares has a potential to grow. That is why we dare to give a recommendation to buy Nike's shares in Q1 2018.

References

Nike Revenue. Retrieved October 03, 2017, from https://ycharts.com/store/nike

Nike's EPS. Retrieved October 03, 2017, from https://ycharts.com/store/nike

Nike's operating margin. Retrieved October 03, 2017, from https://ycharts.com/store/nike

Nike's Dividends. Retrieved October 03, 2017, from https://ycharts.com/store/nike

Nike's Payout Ratio. Retrieved October 03, 2017, from https://ycharts.com/store/nike

WACC. Retrieved November 12, 2017, from https://www.gurufocus.com/term/wacc/NKE/WACC/Nike+Inc.

Nike Inc. Retrieved November 12, 2017, from http://search.morningstar.com/sitesearch/search.aspx?s=o&q=Nike

Gordon Growth Model. Retrieved December 24, 2017, from https://www.investopedia.com/terms/g/gordon-growthmodel.asp

Оценка бизнеса корпорации Nike

Елена Мирошина (Силантьева)[1], Егор Романов[2]

В этой статье мы сосредоточили внимание на способности анализировать рынок и проводить комплексную оценку компании, что является одной из наиболее важных задач перед финансистами. Мы попытались понять, переоценена ли компания Nike, Incorporated или нет. В работе мы провели анализ результатов деятельности Nike, Inc. посредством дискретного денежного потока (DCF) в качестве метода и анализа финансовой отчетности.

Ключевые слова: инвестиционная оценка; DCF; анализ финансовой отчетности; корпорация Найк

JEL classification: D22, F23, G32

[1] доцент, Департамент мировой экономики и мировых финансов, Финансовый университет, Москва; EASila@yandex.ru

[2] Международный финансовый факультет, Финансовый университет, Москва; egor-romanov98@mail.ru

Review of Business and Economics Studies
2018, Vol. 6, No. 1, 83–92
DOI: 10.26794/2308-944X-2018-6-1-83–92

"Uniform Subsidy" and New Trends in Financing of Agricultural Insurance in Russian Federation

L. Yu. Piterskaya[1], N.A. Tlisheva[2], A.V. Piterskaya[3]
Trubilin Kuban State Agrarian University

Abstract

The relevance of the topic is explained by the incomplete content of economic models that form the basis for the distribution of state support in the regions, the discrepancy between the expected reactions and real ones, weakness of the results of the implementation of state programs to support the subjects of the agro-industrial complex (hereinafter — agribusiness) as concerns target indicators, and existing mismatch of selected factor indicators and effectiveness ones. In the article, we have proven the absence of a direct correlation between the method of estimating the number of subsidies aimed at supporting the achievement of the target indicators of regional programs and specific economic content of insurance, with the help of factorial (regression) and retrospective analysis. Therefore, the results of this study can serve as a basis for changing the existing model estimates the number of subsidies aimed at supporting the achievement of target indicators of regional programmes—at least in the insurance industry. In the future, it should allow increasing the efficiency of budget financing of activities related to agricultural insurance.

Keywords: "uniform subsidy"; agricultural insurance; regional program of support; livestock; crop production; federal budget; estimated budgetary efficiency
JEL classification: Q14, Q18

[1]Ludmila Yurievna Piterskaya, D. Sci. (Economics), Professor, Head of Department of monetary circulation and credit, Trubilin Kuban State Agrarian University, Krasnodar, Russia; lpiterskaya@yandex.ru
[2]Nafset Adamovna Tlisheva, PhD (Economics), Associate Professor, Trubilin Kuban State Agrarian University, Krasnodar Russia; cellula@mail.ru
[3]Anastasiya Vladimirovna Piterskaya, master student, Trubilin Kuban State Agrarian University, Krasnodar, Russia; apiterskaya@yandex.ru

When the concept of "uniform subsidy" was introduced it provided a change of the place and role of agricultural insurance subsidization in the concept of agricultural regulation (Burlakova, 2016) (Fig. 1).

When we talk about the role of agricultural insurance in concept of agricultural regulation in the past (in The State Program of Development in Agriculture and for Regulation of Markets of Agricultural Products, Resources, and Food—then we say just State Program) we can notice that agricultural insurance was isolated from two activities: support of economically significant regional programs in Russia—in the field of livestock and crop production (Macht, Makenova, & Karpova, 2017).

However, these activities were realized together with activities "Risk management in sub-sectors of crop production" and "Risk management in sub-sectors of livestock" in the area of subroutines of development of subindustries

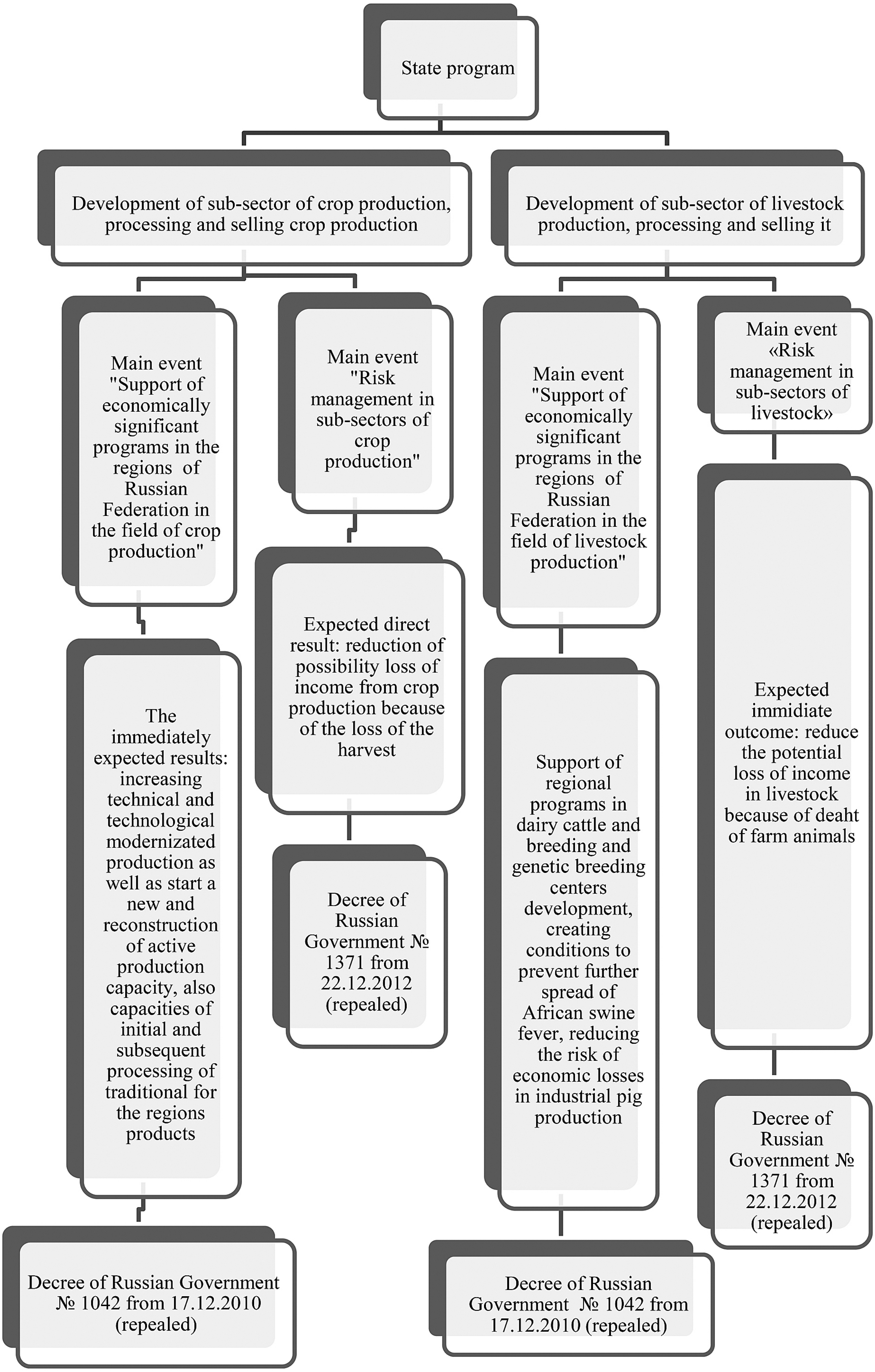


Figure 1. A retrospective look at the state support of agriculture in the field of programs aimed at the development of production of crops and livestock.

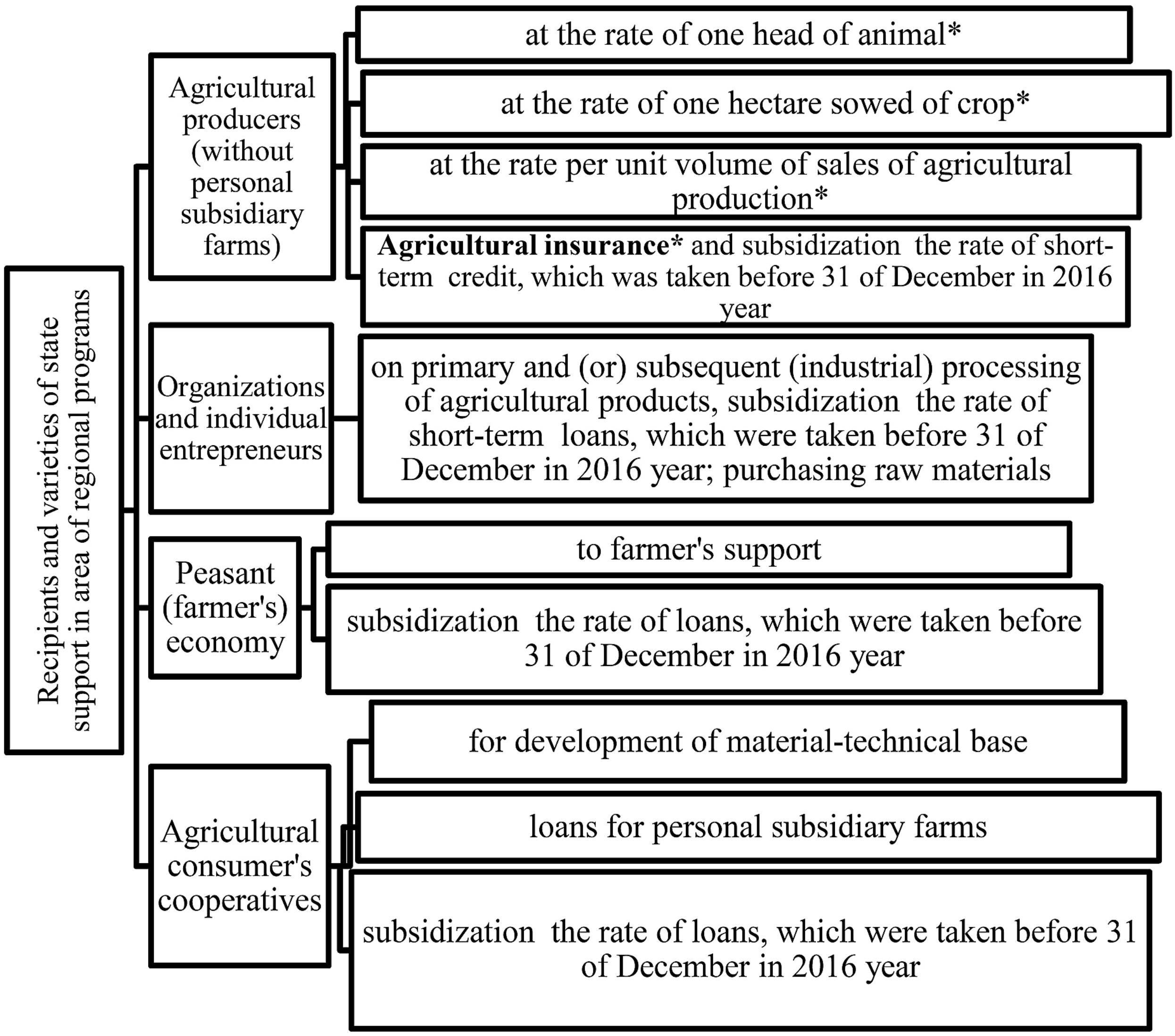


Figure 2. Recipients and types of state support in the field of "regional programs" ("uniform subsidy").*

*Adadimova L. Yu. and others point out to the unity of independent support and agro-insurance, which together belongs to so-called "yellow basket" of World Trade Organization (Adadimova & Polulyakh, 2015; Aleksandrova & Dolbilova, 2015).

in crop production and of subindustries in livestock as uniting beginning.

State support of agro-insurance as well as ensuring food security, saving in the future the traditional for the regions agricultural products, small business development became a part of activity system to achieve targets of regional programs of agriculture development (part of "uniform subsidy"). It is regulating by rules of provision and distribution of subsidies from the federal budget to regional budgets, where regions can determine the direction of spending on their own (Fig. 2 and Fig. 3).

There were two concepts of "regional" subsidies:

1) direct support of regions within such main events as support of economically significant programs in the regions of Russian Federation in crop production and livestock production (without agro-insurance). This system (Fig. 1) was valid in 2010–2014 and described in Decree of Russian Government No. 1042 of Dec. 17, 2010. There were the following guarantees of this type of support:

a) agricultural producers— for the organization of production and processing of agricultural products;

b) organizations engaged in the production of amino acids for animal feed;

c) organizations engaged in the production of wines with protected geographical indica-

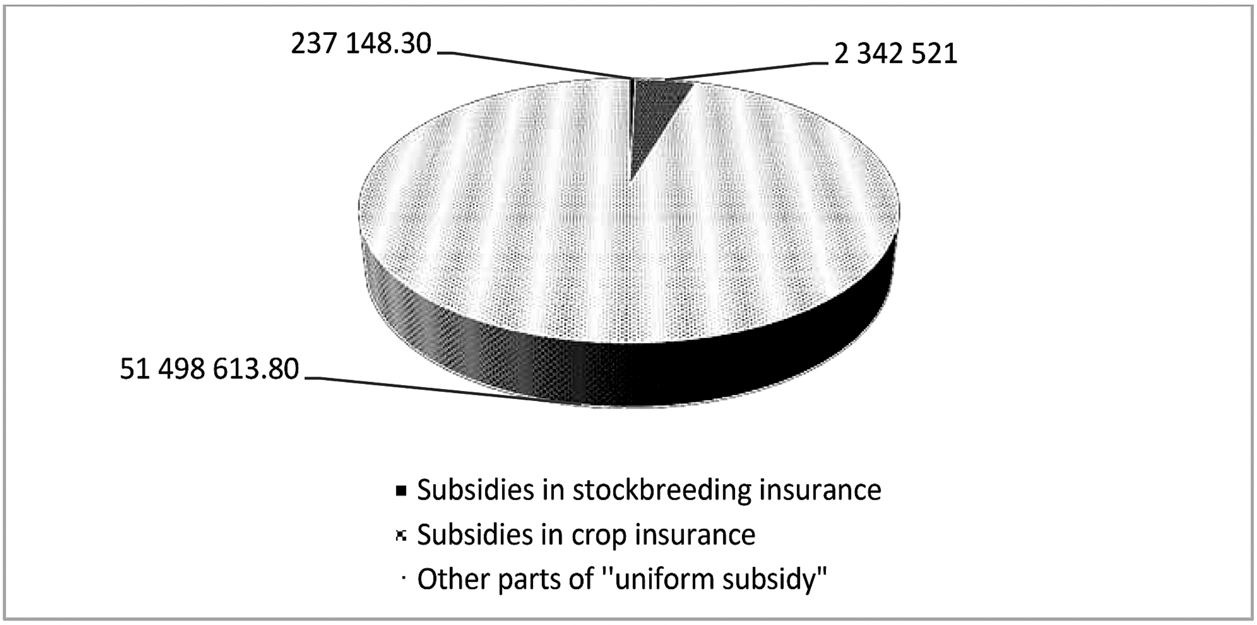


Figure 3. Amount of "uniform subsidy", thousand rubles, 2016 year.

*State support system in agricultural insurance in Russian Federation is organized like in Turkey (insurance pool Tarsim) and Spain (insurance pool Agro Seguro). There is The National Union of Agricultural Insurers in the Russian Federation (Korneev & Kapitonov, 2017).

Sources: 1. State Program — the total amount of financing of "uniform" subsidy from the Federal budget ("Help for achievement targets of realization regional programs of development in agriculture"). Retrieved Dec. 18, 2017, from http://programs.gov.ru/Portal/programs/subActionsList?gpId=27&pgpId=E7F34F65–73A8–48D0-BE11-AA268FFE8B54

tion and protected appellation of origin in the Republic of Crimea and Sevastopol—on a bookmark and care of vineyards, including stubbing retired from service of old vineyards.

In total, in 2016 the largest amount of the budget was accounted for funding in the subroutine of the development of the subsectors of crop production (28%) of the overall appropriation. In 2014, for the development of livestock sub-sector, it was directed 33.87% of funds (Bogoviz et al., 2017).

2) "regional" subsidy as a part of the main program "Help for achieving the targets of realization of regional programs of development in agriculture" ("uniform subsidy"). Second, include these elements which are shown in Fig. 2, also — agricultural insurance. The third is showing actual meaning of "regional" subsidy.

2. From FSBI "Federal Agency of state support of agriculture" of Ministry of Agriculture of Russia. Retrieved Dec. 18, 2017, from http://fagps.ru/sites/default/files/merged%20%281%29_0.pdf

There are differences in targets, methods, and principles of state support and unification and classifications of support (it becomes consolidated today (Sokolova, 2017)).

Discussion about the method

Today, for example, value of subsidies of i^{th} regional budget (W_i), for helping to achievement targets of realization regional programs of development in agriculture is calculated according to formula (1), which directly shows correlation with small business development and indirect correlation with other sides of "uniform subsidy" (agro-insurance with state support, saving a future of traditional agricultural products for the regions):

$$W_i = W \times \frac{\left(V_i + P_i + S_i + K_i\right) / EBS_i}{\sum_{i=1}^{n}\left(V_i + P_i + S_i + K_i\right) / EBS_i}, \qquad (1)$$

where:

W — subsidies providing in federal budget for helping to achievement targets of realization regional programs of development in agriculture in a current financial year;

V_i — portion of i^{th} region in total volume of crop production and stockbreeding and food

Table 1
The discrepancy between the parameters that determine the distribution of regional subsidies and indicators for assessment of effectiveness of regional budgets funds' distribution

Parameters determining the distribution of regional subsidies	**Indicators for assessment of effectiveness of regional budgets funds' distribution**
a) preservation of regions' traditional agricultural products	
a) the number of breeding stock of sheep and goats b) the number of conditional breeding stock of breeding animals, and so on	a) the breeding stock of sheep and goats in agricultural organizations; peasant (farmer) farms, including individual entrepreneurs (thousand units) b) preservation of the conditional tribal breeding stock of farm animals to the level of the previous year (%) c) the realization of breeding young cattle of dairy and beef breeds for the 100 heads of female (heads), and so on
b) agro-insurance with state support and ensuring food security	
a) the size of areas under crops, sown seeds in accordance with the list determined by the Ministry of Agriculture of the Russian Federation b) the amount of acreage under fodder crops in the territory of the Russian Federation, carried to regions of the far North and equated localities c) the size of the area of low productive arable land (pure vapor) constituting not less than 11 percent of the total arable land in the territory of the Russian Federation, carried to regions of the far North and equated localities d) the size of perennial fruit and berry plantations e) the size of the area of vineyards and grape nurseries	a) the gross yield of grain and leguminous crops in farms of all categories (thousand tons) b) the gross harvest of sugar beet in farms of all categories (thousand tons) c) the gross yield of flax fiber and pengawalan in all categories of farms (thousand tons) d) the gross harvest of potatoes in agricultural organizations, peasant (farmer) farms, including individual entrepreneurs (thousand tons) e) the production of livestock and poultry for slaughter in all categories of farms (in live weight) (thousand tones) f) the insured livestock (thousand heads) g) the area of preparing low productive arable land (pure vapor) (thousand hectares) h) the fraction of land area, sown with elite seeds total crop area (%) i) area of perennial plantations (thousand hectares) j) the grape plantations in a mature, fruit-bearing age (thousand hectares) k) the size of the insured cultivated area (thousand hectares)
c) small business development	
a) the number of private (peasant) farms and individual entrepreneurs b) the number of agricultural consumer cooperatives, etc.	a) the number of new permanent jobs created in the peasant (farm), to implement the projects of creation and development of their farms by means of government support (units) b) the growth of agricultural output produced by individual entrepreneurs and peasant (farming) enterprises, which received state funding, to the year preceding the year of grant (percent)
d) all sides, factors of the model (see formula (1))	
a) an average volume of crop production and stockbreeding and food production b) the rest of short-term loans, which were taken before 31 of December in the 2016 year	–

Table 2

Grouping of the regions, which are classified in terms of the level of estimated budgetary efficiency

Border of group of regions in terms of the degree of estimated budgetary sufficiency		Region	Number of regions in the group	Among them: number of region recipients of crops and perennial plantings insurance support	Amount of subsidies, thousand rubles	Sown area, thousand hectares	The amount of subsidies per hectare of sown area, rubles
0.637	0.727	Republic of Tuva, Ivanovskaya oblast, Kostromskaya oblast, Orlovskaya oblast, Tambovskaya oblast, Republic of Karelia, Arkhangelskaya oblast, Pskovskaya oblast, Republic of Adygea, Republic of Kalmykia, Ingushetia, Kabardino-Balkaria, Karachayevo-Cherkessia, North Ossetia, Chechen Republic, Stavropol Kray, Republic of Mari' El, Chuvashia, Kirovskaya oblast, Penzenskaya oblast, Kurganskaya oblast, Republic of Altay, Republic of Buryatia, Republic of Khakassia, Altay Kray, Zabaykalsky Kray, Republic of Sakha (Yakutia), Kamchatskiy Kray, Magadanskaya oblast, and Chukotskiy autonom. district, Republic of Crimea, Sevastopol, Republic of Dagestan, Bryanskaya oblast	34	11	782321.0	13734.2	56.96161
0.727	0.817	Smolenskaya oblast, Volgogradskaya oblast, Rostovskaya oblast, Republic of Bashkortostan, Republic of Mordovia, Saratovskaya, Ulyanovskaya, Kemerovskaya, Primorsky kray, Khabarovsky kray, Jewish autonom. oblast, Ryazanskaya oblast, Vladimirskaya oblast, Voronezhskaya oblast, Omskaya oblast	15	11	512350	23 128	22.15233
0.817	0.907	Belgorodskaya oblast, Kurskaya oblast, Tverskaya oblast, Kaliningradskaya oblast, Novgorodskaya, Astrakhanskaya oblast, Krasnodar Kray, Udmurtia, Permskiy kray, Orenburgskaya oblast, Chelyabinskaya oblast, Irkutskaya, Novosibirskaya, Tomskaya oblasts, Amurskaya oblast, Vologodskaya oblast	16	9	845082.000	15693.075	53.85063
0.907	0.997	Lipetskaya, Tulskaya oblasts, Komi Republic, Murmanskaya oblast, Nizhegorodskaya oblast, Krasnoyarskiy kray	6	3	132205	3306	39.9879
0.997	1.087	Yaroslavskaya oblast, Kalugskaya oblast	2	–	–	–	
1.087	1.177	Nenetskiy autonom. district, Samarskaya oblast, Sverdlovskaya oblast	3	2	107406	2902	37.00723
1.177	1.267	Moscow oblast, Tatarstan	2	2	127625	3626	35.19892
1.267	1.357	Leningradskaya oblast	1	1	4867	240	20.25562
1.357	1.447	Sakhalinskaya oblast	1	1	3304	29	115.472
1.627	1.717	Khanty-Mansiyskiy autonom. district	1	–	–	–	
1.897	1.987	Saint Petersburg	1	–	–	–	
1.987	2.077	Tyumenskaya oblast, Yamalo-Nenetskiy autonomic district	2	–	–	–	
TOTAL			84	40	2515160.0	62658.9	40.14052
IHH groups of specific budgetary provision regions in terms of subsidies					2583.025		
IHH regions in terms of subsidies					916.749		

Note. Herfindahl–Hirschman Index – IHH.

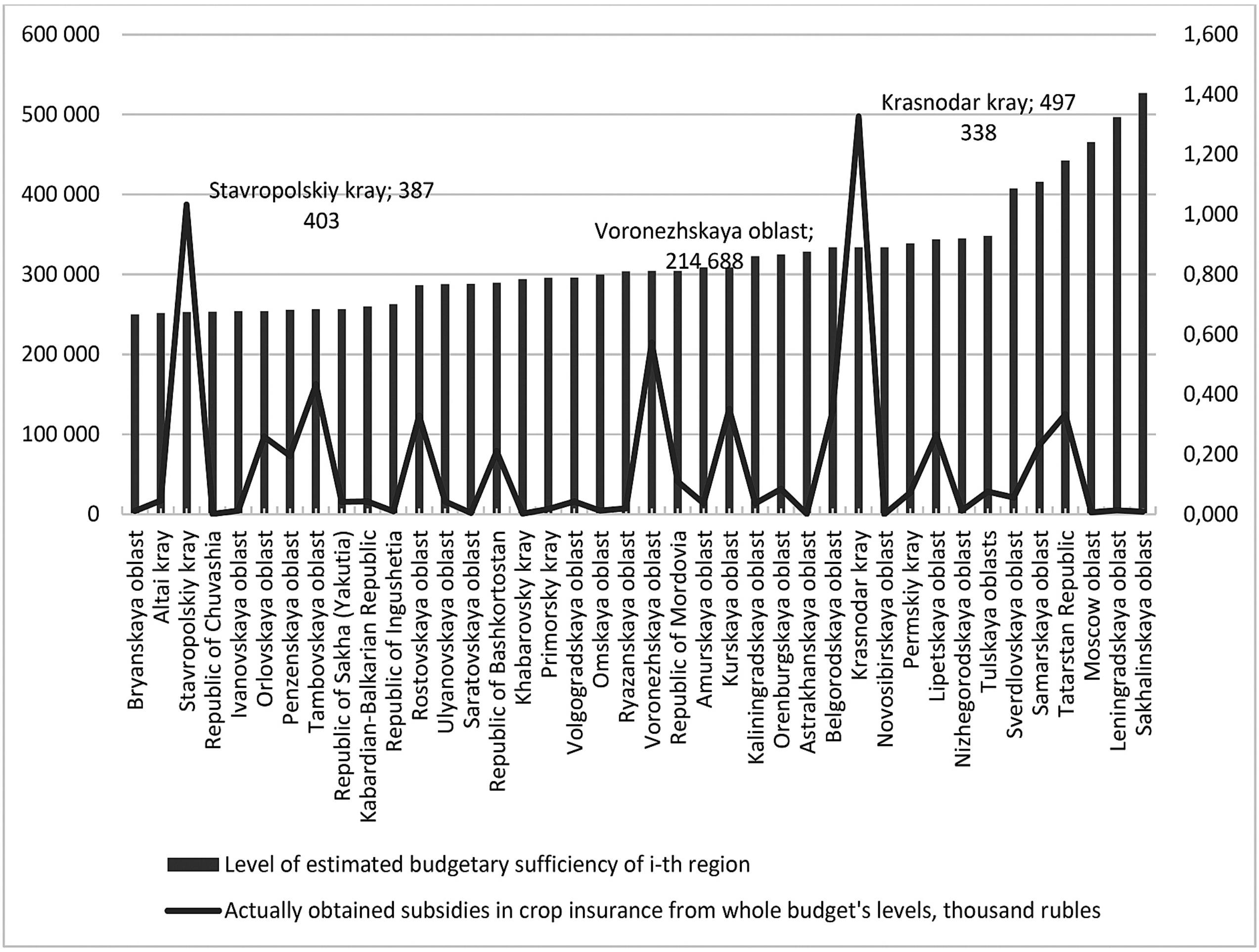


Figure 4. Characteristics of the estimated budgetary effects of subsidies for recipient's regions with crop's agro-insurance.

production and in the balance of outstanding short-term loans, which were taken before 31 of December in the 2016 year;

P_i — share of i^{th} region in the size and growth of the livestock;

S_i — share of i^{th} region in the size of crop area;

K_i — share of i^{th} region in the number of peasants (farmer's) economy, include individual entrepreneurs, and agricultural consumer's cooperatives, volume of peasant (farmer's) economy's and individual entrepreneurs production, and in the balance of outstanding loans to small business;

EBS_i — level of estimated budgetary sufficiency of i^{th} region in a current financial year;

n — the number of the regions.

The volume of the unused financing, allocated for subsidizing of expenses on payment of insurance premiums, fully returning to the federal budget now (Belova & Sannikova, 2017).

We think that our modern system of "uniform subsidy", as a particular case of imperfections, confirms Krugman's justice who blaming modern economists because they are too fascinated by the mathematical elegance of their models, forgetting about the content of economic processes (see http://www.econorus.org/fmean.phtml).

You can see immediately the discrepancy between the parameters that determine the distribution of regional subsidies and indicators for assessing the effectiveness of implementation of expenses of regional budgets (only for agro-insurance) at Table 1.

Besides the methodological issues we see problems of failure big share of insured crop area in the size of crop area (in 2016–5%, in 2015–10.9%), state planes the volume of insured crop area in the level 4067.7 thousand hectares, whole size of crop area for crop in 2017 79993.038 thousand hectares, so we can wait share of insured crop area in 5.1%; some regions didn't receive support in 2016 in insurance but they need it, for example, Kalmyk republic, Republic of Crimea and Sevastopol, Republic of Adygea. Kalmyk republic is situated on the territory of the agricultural zone with a probability of severe droughts ⩾50% in the period from May to August.

In 2016 there are 40 regions, which was subsidized (Fig. 4). By the level of estimated budgetary sufficiency big groups of regions are relatively homogeneous (Table 2). In those groups also we see compliance with a number of regions in each group of the level of estimated budgetary sufficiency and numbers of regions in a group of regions-recipients of subsidies.

But the dependence is not revealed between "the actually obtained subsidies in crop insurance from whole budget's levels" (y) and "the level of estimated budgetary sufficiency of i^{th} region" (x):

$$y = -489333x+104579\ (R^2 = 0.0074);$$

or

$$y = -40940\ln(x) + 55529\ (R^2 = 0.0058);$$

or

$$y = -236480x^2 + 415623x - 111853\ (R^2 = 0.0159),$$ or other models.

IHH consolidated groups of the regions of a specific budgetary provision in the volume of subsidies are very high, which serves as an indicator of state preferences in the financing of regions with a certain level of fiscal capacity, which in turn has a negative impact on the agricultural insurance system. Even in terms of the need to save the acreage of fodder crops in agricultural organizations, peasant (farmer's) farms, including individual entrepreneurs, in the far North and equivalent areas, the situation when one enterprise in Sakhalinskaya oblast in 2016, receives support amounting to 3304 thousand rubles, and in the 11 regions with budget sufficiency from 0.727 to 0.817, to get 22 rubles per 1 ha of sown area cannot be considered as normal.

Conclusions

In closing, let's make four important points:

- it is necessary to change approaches to the formation of methods of determining funding of the "uniform subsidies", or absolutely reject the concept of a "uniform subsidy" in favor of financing of food security, of small business, of priority of traditional industries and agricultural insurance;
- it is necessary to establish a system of indicators for the distribution of grants between budgets of the regions of Russian Federation in agro-insurance ("the parameters determining the distribution of regional subsidies" for insurance in table 1 in this article), follows directly from indicators of the use of subsidies in agro-insurance (column 2 of table 1 in this article);
- it is necessary to complement the performance indicators of the use of subsidies in agro-insurance, basing at network by All-Russian Research Institute of agricultural meteorology observations of air temperature on the territory of Russia, anomalies of average air temperature during the vegetation period of spring cereals from date of germination to date of harvest and others parameters, for example, results of monitoring of agro-climatic conditions of yield formation of crops;
- it is necessary to base the calculation of the W_i from formula (1) not so much on the budget provision, when agricultural insurance is object of analyses, there are many other indicators of the variability of agro-climatic growing conditions of crops (Trubilin et al., 2016); or rental conditions (Klishina & Uglickih, 2017), or climate indices indicative for the insurance case — the lack of rainfall in the area for a certain number of days, that is a kind of "futures" on the weather conditions (Vanyushina, 2014), that can replace or add to an estimated budgetary sufficiency.

References

Adadimova, L. Yu., & Polulyakh, Yu. G. (2015). Ustoichivyi rost i strakhovanie riskov agrarnogo proizvodstva kak osnova prodovol'stvennoi bezopasnosti [Sustainable growth and insurance of risks of agricultural production as the basis of food security]. *Ostrovskie chtenija*, *1*, pp. 141–144. Retrieved from https://elibrary.ru/item.asp?id=24867757.

Aleksandrova, L. A., & Dolbilova, E. A. (2015). Gosudarstvennaya podderzhka sel'skogo khozyaistva v Rossii: novye formy i prioritety [State support of agriculture in Russia: new forms and priorities]. *Agrarnyj nauchnyj zhurnal, 3*, pp. 71–76. Retrieved from https://elibrary.ru/download/elibrary_23173612_58361489.pdf.

Belova, E. V., & Sannikova, M. O. (2017). Retrospektivnyi analiz i perspektivy razvitiya sistemy subsidirovaniya dogovorov sel'skokhozyaistvennogo strakhovaniya [Retrospective analysis and prospects of development of the system of subsidizing of agricultural insurance contracts]. *Agrarnyj nauchnyj zhurnal, 9*, pp. 75–82. Retrieved from https://elibrary.ru/download/elibrary_30057164_69070413.pdf.

Bogoviz, A. V., Rykova, N. N., Shkodinski, S. V., & Gubanov, R. S. (2017). Mery gospodderzhki regional'nykh programm razvitiya agropromyshlennogo kompleksa Severo-Kavkazskogo i Yuzhnogo federal'nykh okrugov [State support of regional programs of development of agroindustrial complex of the North Caucasus and Southern Federal districts]. *Ekonomika sel'skohozjajstvennyh i pererabatyvajushhih predprijatij*, *9*, pp. 13–18.

Burlakova, Ekaterina. (2016, August 19). The Ministry of agriculture intends to combine subsidies, reducing their number. The current scheme does not suit farmers. *Vedomosti.* Retrieved from https://www.vedomosti.ru/business/articles/2016/08/19/653720-minselhoz-nameren-obedinit-subsidii-sokrativ-chislo.

Klishina, Y., & Uglickih, O. (2017). Mezhdunarodnyi opyt strakhovaniya sel'skokhozyaistvennykh riskov [International experience in agricultural insurance]. *Ekonomika sel'skogo hozjajstva Rossii, 8*, pp. 91–98.

Korneev, A. F., & Kapitonov, A. A. (2017). Novye mery stimulirovaniya realizatsii gosudarstvennoi programmy razvitiya sel'skogo khozyaistva [New measures to boost the implementation of the state program of development of agriculture]. *Ekonomika sel'skogo hozjajstva Rossii*, *6–1*(60), pp. 29–32.

Macht, V. A., Makenova, S. K., & Karpova, O. A. (2017). Analiz sushchestvuyushchei metodiki klassifikatsii zemel' [Analysis of the existing methods for the classification of land]. *Vestnik Voronezhskogo gosudarstvennogo agrarnogo universiteta*, *1*, pp. 253–258. Retrieved from https://elibrary.ru/item.asp?id=29282583.

Sokolova, I. A. (2017). Konsolidatsiya gosudarstvennoi podderzhki i sud'ba agrostrakhovaniya [Consolidation of state support and the fate of crop insurance] *Materialy mezhdunarodnoj nauchno-prakticheskoj konferencii* "*Nauka i obrazovanie: opyt, problemy, perspektivy razvitija*" [Materials of the international scientific-practical conference "Science and Education: Experience, Problems, Prospects"]. Krasnoyarsk. Krasnoyarsk state agrarian university, pp. 289–291.

Trubilin A. I., Piterskaya L. Yu., Voroshilova I. V., Tlisheva, N. A., & Baranovskaya, T. P. (2016). The anti-crisis strategy of state regulation reproductive processes in agriculture. *International Journal of Economic Research,* 13(9), pp. 3885–3909.

Vanyushina, O. I. (2014). Metodicheskie podkhody k formirovaniyu sistemy strakhovaniya v agrarnom sektore ekonomiki [Methodical approaches to formation of system of insurance in agrarian sector of economy]. *Vestnik Rjazanskogo gosudarstvennogo agrotehnologicheskogo universiteta im. P. A. Kostycheva, 3*(23), pp. 62–65.

Acknowledgements

The study was performed with the financial support of RFFR and Administration of Krasnodarskiy Kray as a part of the scientific project No. 17–12–23032 a(r).

«Единая субсидия» и новые веяния в субсидировании агрострахования в Российской Федерации

Людмила Питерская[1], Нафсэт Тлишева[2], Анастасия Питерская[3]

Актуальность темы объясняется неполнотой содержания экономических моделей, составляющих основу распределения государственной поддержки в регионах, несоответствием ожидаемых реакций, результатов реализации государственных программ поддержки субъектов агропромышленного комплекса (далее — АПК) целевым ориентирам этих моделей и стимулам, направленным на объект регулирования (рассогласованностью избранных факторных показателей и результативных).

В статье обосновывается отсутствие прямой взаимосвязи метода оценки объема субсидий, направленных на поддержку достижения целевых показателей региональных программ, с содержательной экономической спецификой страхования, доказанное факторным (регрессионным) анализом, ретроспективным анализом. Методы, используемые в статье: обобщение, абстрагирование, конкретизация, мысленный эксперимент, а также графический метод.
Результаты настоящего исследования могут послужить основой для изменения существующей модели оценки объема субсидий, направленных на поддержку достижения целевых показателей региональных программ — по крайней мере, в сфере страхования, что в перспективе может способствовать увеличению бюджетной эффективности финансирования мероприятий, связанных с сельскохозяйственным страхованием.
Ключевые слова: «единая субсидия»; агрострахование; региональная программа поддержки; животноводство; растениеводство; федеральный бюджет; расчетная бюджетная эффективность
JEL classification: Q14, Q18

Благодарность

Исследование выполнено при финансовой поддержке РФФИ и Администрации Краснодарского края в рамках научного проекта № 17–12–23032 а(р).

[1] Доктор экономических наук, профессор, заведующий кафедрой денежного обращения и кредита, Кубанский государственный аграрный университет имени И.Т. Трубилина, Краснодар, Россия; lpiterskaya@yandex.ru
[2] Кандидат экономических наук, доцент, Кубанский государственный аграрный университет имени И.Т. Трубилина, Краснодар, Россия; cellula@mail.ru
[3] Магистрантка, Кубанский государственный аграрный университет имени И.Т. Трубилина, Краснодар, Россия; apiterskaya@yandex.ru