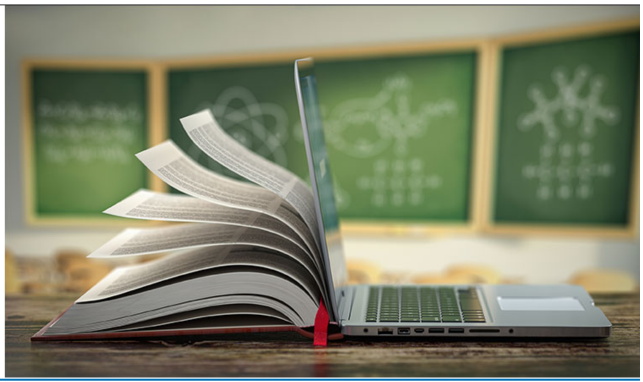
PAPER • OPEN ACCESS

Sustainable Development Indicators in Benchmarks of Russia's Regional Policy

To cite this article: Alexander N Tyrsin and Elena V Vasilyeva 2020 *IOP Conf. Ser.: Earth Environ. Sci.* **534** 012003



Sustainable Development Indicators in Benchmarks of Russia's Regional Policy

Alexander N Tyrsin[1] and Elena V Vasilyeva[2]

[1] Chair of the Applied Mathematics Department, UrFU, Ekaterinburg, Russia
[2] Senior Research Associate, Institute of Economics, Ural Branch of the Russian Academy of Sciences, Ekaterinburg, Russia

E-mail: elvitvas@ya.ru

Abstract. We have considered the problem concerning the substantiation of sustainable development indicators and the improvement of the standard of living of the population in Russian regions. To quantify the standard of living of the population, we have used a common indicator such as the Human Development Index. The authors present an approach to managing the region in the form of optimization problems aimed at improving the standard of living of the region's population by increasing the probability of its classification as a region with a higher standard of living. The optimization problem is solved based on the identified relationship between the standard of living of the region's population and its socioeconomic indicators.

1. Introduction

Determining practical measurable indicators of the sustainable development concept is of essence for its implementation. There is a wide range of approaches and indicators offered by leading international organizations (UN, World Bank, OECD, European Commission, etc.). These indicators are used in the program-based approach to territory management as markers determining the system sustainability.

Russian regions are developing their systems of sustainable development indicators using the current database of the official Russian statistics which, according to S.N. Bobylev [1], makes it much easier to obtain necessary information and provides an adequate assessment of the regional progress in sustainable development. Since 2012, the general Russian regional development strategy has been determined by the so called "May Decrees" that contain clear benchmarks in such areas as education, science, healthcare, economy, demography and utilities. The goal of this research is to understand whether the target indicators of the regional policy are well justified, and to what extent their accomplishment will promote the region's sustainable development and improve the standard of living of its population.

2. Building a logistic regression by the criterion of maximum likelihood

Let us take a training sample by a set of input features $X_1, X_2, \ldots, X_m$

$$(\boldsymbol{x}_i,\ y_i),\ i = 1,\ 2,\ \ldots,\ n, \tag{1}$$

where $\mathbf{x}_i$ is the vector of values of the i-th object $\boldsymbol{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1m} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{n1} & \cdots & x_{nm} \end{pmatrix}$; $\boldsymbol{y} = \begin{pmatrix} y_1 \\ \cdots \\ y_n \end{pmatrix}$, $y_i \in \{-1;\ 1\}$ is the binary variable indicating the categorization of the i-th object into a relevant class, for example, the first class, if $y_i = -1$, or the second class, if $y_i = 1$; *m* is the number of features of each object; *n* is the number of observations. The classification is performed using the function

$$h(\boldsymbol{x}) = \frac{1}{1+exp\{-\boldsymbol{b}^T\boldsymbol{x}\}}, \quad (2)$$

taking values in the range of (0; 1). The threshold value is $h(\mathbf{x}) = 0{,}5$. The function $\delta(t) = \dfrac{1}{1+e^{-t}}$ is called the logistic function. The vector $\mathbf{b}^T = (b_0\ b_1\ ...\ b_m)$ in (2) sets a separating linear border described with the hyperplane equation

$$\Pi: \quad \boldsymbol{b}^T\boldsymbol{x} = 0. \quad (3)$$

Let us introduce the function

$$W(\boldsymbol{x}) = \boldsymbol{b}^T\boldsymbol{x}. \quad (4)$$

Let us set the D_1 area of possible **x** values for the first class as $D_1 = \{\mathbf{x} : W(\mathbf{x}) < 0\}$, and for the second class as $D_2 = \{\boldsymbol{x}: W(\boldsymbol{x}) > 0\}$. Then $\forall\boldsymbol{x} \in D_1\ h(\boldsymbol{x}) < 0.5$ and $\forall\boldsymbol{x} \in D_2\ h(\boldsymbol{x}) > 0.5$. If **x** belongs to the hyperplane (3), then $h(\mathbf{x}) = 0{,}5$. Consequently, for the arbitrary observation $\mathbf{x}^*$, the probability of its placement into the first class is equal to $P(\boldsymbol{x}^* \in D_1) = 1 - h(\boldsymbol{x}^*)$, and to the second class – $P(\mathbf{x}^* \in D_2) = h(\mathbf{x}^*)$. It is established [2] that the maximization of the likehood logarithm is equivalent to the minimization

$$Q(\boldsymbol{b}) = \sum_{i=1}^{n} ln\left(1 + e^{-y_i\boldsymbol{b}^T\boldsymbol{x}_i}\right) \to \min_{\boldsymbol{b}\in\boldsymbol{R}^m}. \quad (5)$$

To assess the coefficient vector **b**, iteration descent algorithms are used to solve extremum problems, e.g. the Newton–Raphson iteration algorithm is offered [2–3]. It involves the following. As a zero approximation, the classification problem can be accomplished using the multivariable linear regression method $\mathbf{b}^{(0)} = (\mathbf{X}^T\mathbf{X})^{-1}(\mathbf{X}^T\mathbf{y})$. Then the iteration process starts, on the *k*-th step of which the coefficient vector **b**is specified$^{(k)}$: $\mathbf{b}^{(k)} = \mathbf{b}^{(k-1)} - h_k(Q''(\mathbf{b}^{(k-1)}))^{-1}Q'(\mathbf{b}^{(k-1)})$, where $Q'(\mathbf{b}^{(k)})$ is a vector of first derivatives (gradient) of the $Q(\mathbf{b})$ functional in the $\mathbf{b}^{(k)}$point, $Q''(\mathbf{b}^{(k)})$ is a matrix of second derivatives (hessian) of the $Q(\mathbf{b})$ functional in the $\mathbf{b}^{(k)}$point,andh_k is a step value that can be equivalent to 1, although its selection at each step can increase the convergence rate. The Newton–Raphson method is described in a number of textbooks on optimization methods [4]. However, the minimization problem (5) is not generally accomplished as, with a correct classification of all precedents, it has the lower bound at infinity equal to zero [5]. The quality of the built separating the border with a formal descent will not be related to the value of the target function $Q(\mathbf{b})$ any more. The growth in the component of the **b** vector results in the unconstrained growth of values – $y_i\mathbf{b}^T\mathbf{x}_i$ that can take very large values, which causes inherent computational errors or even stoppage of the algorithm due to memory overflow [6]. For this reason, the descent strategy needs to be adjusted. Gradient methods are intended for a descent to a local minimum. As the target function has a lower bound at infinity and there can be points, in which the gradient is close to zero, we deem it reasonable to proceed to zero-order methods. In [5], the authors offer an iteration descent mechanism. It comprises a sequence of zero-order minimization problems based on random searching. At every *k*-th step, the length of the coefficient vector $\mathbf{b}^{(k)}$ is fixed. It is gradually increased until the target function reaches the required value.

3. Logistic regression as a mathematical model of relationship between the human development index and the socioeconomic indicators of the region

To characterize the standard of living of the region's population, we use the Human Development Index (HDI). Let us consider the values of the informative socioeconomic indicators of the**i**-th region of Russia as the vector $_{x}i$. We have formed two classes of regions: the first class comprises regions with a low HDI

(HDI < D), and the second class includes regions with a high HDI (HDI ≥ D), where D is a threshold value. For the first D_1 class, let the value of the binary variable membership be $y = -1$, and for the second class D_2, let it be $y = 1$. Let us consider a region with a set of actual values of its socioeconomic indicators as the vector $\mathbf{x}^*$. According to (3), we are assessing $W(\mathbf{x}^*)$. If $W(\mathbf{x}^*) < 0$, the region will fall into the first class with the probability $1 - h(\mathbf{x}^*)$. With $W(\mathbf{x}^*) > 0$, the region will fall into the second class with the probability $h(\mathbf{x}^*)$. If $W(\mathbf{x}^*) = 0$, the case is uncertain: the region can be placed both into the first class and into the second class with an equal probability of 0.5. Thus, a probabilistic assessment of the standard of living of the region's population can be made. The region can be classified among regions with a high standard of living with the probability $h(\mathbf{x}^*)$, and among regions with a low standard of living with the probability $1 - h(\mathbf{x}^*)$. Therefore, the probability $P(\mathbf{x} \in D_2) = h(\mathbf{x})$ can be used as the value of the target function: the higher its value, the more likely it is that the region's population will be deemed as having a high standard of living.

The features of the object in question are determined by the location of the point $\mathbf{x}^*$ relative to the hyperplane (3). They can be numerically described, for example, using the distance between the point $\mathbf{x}^*$ and the hyperplane (3) $d = \left|\mathbf{b}^T \mathbf{x}^*\right| \Big/ \left(\sum_{j=1}^{m} b_j^2\right)^{1/2}$ and the gradient $\operatorname{grad} h(\mathbf{x}^*)$ of the function $h(\mathbf{x})$ in the point $\mathbf{x}^*$ $grad\, h\,(\boldsymbol{x}^*) = \left(\frac{\partial h(\boldsymbol{x}^*)}{\partial x_1}, \ \dots \ , \ \frac{\partial h(\boldsymbol{x}^*)}{\partial x_m}\right) = q(\boldsymbol{x}^*)(b_1, \ \dots \ , \ b_m), q(\boldsymbol{x}^*) = \frac{exp\{-\boldsymbol{b}^T\boldsymbol{x}^*\}}{(1+exp\{-\boldsymbol{b}^T\boldsymbol{x}^*\})^2}$.

Thus, $\operatorname{grad} h(\mathbf{x}^*)$ is always orthogonal to the hyperplane (3), its direction is set by the vector $\mathbf{b}$, and its length – by the value $q(\mathbf{x}^*)$. As the coefficient vector sets the hyperplane (3) equation with the accuracy up to the multiplier, it is expedient to normalize it for definiteness purposes when solving practical problems, i.e. $\mathbf{b}^0 = \mathbf{b}/|\mathbf{b}|$, $\left|\mathbf{b}^0\right| = \sum_{j=1}^{m}(b_j^0)^2 = 1$.

Let us assume that $Z(\mathbf{x}) = \dfrac{h(\mathbf{x})}{1 - h(\mathbf{x})}$, i.e. $\ln Z(\mathbf{x}) = b_0 + \sum_{j=1}^{m} b_j x_j$, $b_k = \dfrac{d_{x_k} Z(\mathbf{x})}{Z(\mathbf{x}) dx_k}$, $k = 1, \dots, m$.

By equating the differential $dx_k = \Delta x_k = 1$ and multiplying both parts of the latter formula by 100%, we receive $b_k \cdot 100\% \approx \dfrac{\Delta_{x_k} Z(\mathbf{x})}{Z(\mathbf{x})} \cdot 100\%$. As $Z(\mathbf{x}) = \dfrac{P(\mathbf{x} \in D_2)}{P(\mathbf{x} \in D_1)}$, the coefficient b_k with the variable X_k shows the percentage, by which the ratio of the probability of the region's classification as one with a high standard of living (D_2) to the probability of classifying the region among regions with a low standard of living (D_1) will change, if X_k increases by one unit, with the fixed values of other variables.

The derivative of the function $Z(\mathbf{x})$ for the variable X_k is directly proportionate to the coefficient b_k. The growth (reduction) of the function $Z(\mathbf{x})$ is equivalent to the growth (reduction) of $h(\mathbf{x})$. Accordingly, to increase the value $h(\mathbf{x})$, with positive values b_k the increments Δx_k must also be positive, or negative, if the values are negative.

One of the main issues related to the use of logistic regression is the ability to establish a statistically reliable HDI-based classification of regions. Let us consider the classification of regions in two classes [7]: with a low HDI (y=–1) and with a high HDI (y=1). Based on the statistics for 2013, two classes of precedents (training samples) were formed. The first class of regions with a low HDI (maximum 0.84) included 28 subjects of the Russian Federation. The second class of regions with the HDI of over 0.85 included 31 subjects of the Russian Federation. To improve the reliability of results, we removed from the total array of statistical data 24 constituent entities of the Russian Federation with very low and very high HDIs, as well as those whose HDI was close to the average value, and those with data gaps. As a result, we obtained a system of statistically significant indicators which enabled us to accurately classify the observations for the training samples. The coefficients of the separating hyperplane (3) equation built on the basis of the training sample are given in Table 1. With variables X_1,

X_6, X_7, X_9, X_{10}, X_{12}, X_{13}, the coefficients of logistic regression turned out to be negative. That is, with fixed values of all other variables, for the improvement in the region's standard of living [the probability $h(\mathbf{x})$], the values of these variables should be decreased. For other variables, the probability $h(\mathbf{x})$ can be increased by increasing their values (with fixed values of other variables). The highest positive value of the coefficient is obtained, if the value of investments in fixed capital per capita is (X_4) [7]. This is quite reasonable and consistent with Keynesian theory. One of the main requirements for the model is its adequacy to the described system or value. The coefficients of the hyperplane (2) equation were built using the data for 2013 (Table 1).

Table 1. Indicators and coefficients of the separating hyperplate (3) equation for the classification of constituent entities of the Russian Federation by the Human Development Index.

Indicator	Symbol	Coefficient (b_j)
Constant term, 1	X_0	0.072
Total birth rate, units	X_1	–2.381
Life expectancy at birth, years/100	X_2	0.807
Number of mortgage loans granted by credit institutions to resident individuals, units per 1,000 people	X_3	0.879
Investments in fixed capital per capita, rubles per 100 thousand people, rubles/10^5	X_4	10.004
Gross regional product per capita, million rubles per individual	X_5	3.974
Ratio of investments in fixed capital to gross regional product, units * 10	X_6	–1.911
Mortality from neoplasms, number of deaths per 1,000 people	X_7	–1.149
Ratio of the average wage earned by junior medical staff to the average wage in a constituent entity of the Russian Federation, units * 10	X_8	0.165
Ratio of the average wage earned by mid-level medical staff to the average wage in a constituent entity of the Russian Federation, units	X_9	–1.770
Total unemployment rate, %/10	X_{10}	–2.739
Availability of physicians, individual per 1,000 people	X_{11}	0.688
Consolidated budget of a constituent entity of the Russian Federation and territorial state extra-budgetary fund for housing and utilities, thousand rubles per 10 people	X_{12}	–4.223
Consolidated budget of a constituent entity of the Russian Federation and territorial state extra-budgetary fund for health care, thousand rubles per 10 people	X_{13}	–0.870

We examined the adequacy of the diagnostic model (1) based on the statistical data released by Rosstat for the next year of 2014, using the case of the subjects of the Ural Federal District. Table 2 presents the actual HDI values resulting from the assessment of the $h(\mathbf{x})$ probabilities over 2013–2014. It should be noted that in 2013 the $h(\mathbf{x})$ probabilities were based on training samples; and in 2014 – on test statistics available under the formula (2). When establishing the quantitative values of attributes X_4, X_5, X_{12}, and X_{13} for 2014, we recalculated the monetary indicators taking into account the inflation estimated through the consumer price index.

Table 2. The actual values of the HDI and the assessment of the $h(\mathbf{x})$ probabilities of classifying the subjects of the Ural Federal District into the group of regions with a high standard of living.

Indicator	Year	Kurgan Region	Sverdlovsk Region	Tyumen Region	Chelyabinsk Region
HDI	2013	0.829	0.868	0.901	0.848
	2014	0.831	0.873	0.903	0.857
$h(\mathbf{x})$	2013	0.005	0.851	1.000	0.805
	2014	0.015	0.936	1.000	0.884

The results presented in Table 2 indicate the adequacy of the model. The increase in the HDI values for the Kurgan, Sverdlovsk and Chelyabinsk Regions correlates with the increase in the $h(\mathbf{x})$ probabilities for all the three regions. The Tyumen Region has the maximum value of its HDI in both cases, which corresponds to the maximum probability $h(\mathbf{x})$ that is practically equal to 1.

4. Optimization models based on logistic regression

The earlier publication [7] describes the approach to the region management in the form of optimization problems intended to improve the standard of living of the region's population by increasing the $h(\mathbf{x})$

probability of its classification as a region with a higher standard of living. The higher standard of living is achieved due to an optimum change in the vector of socioeconomic indicators **x**. Let us consider these problems. The first one involves maximization of the $h(\mathbf{x})$ probability subject to restrictions on changes in the indicators:

$$\begin{cases} h(\mathbf{x}) \to \max, \\ x_j = x_j^0 + \Delta_j,\ j = 1, \ldots, m, \\ \Delta_j \in G_j,\ j = 1, \ldots, m, \end{cases} \tag{6}$$

where Δ_j is the change in the j-th component, G_j is the area of admissible values of the change in the j-th component, and $\mathbf{x}^0$ is the vector of current values of the region's indicators.

The problem (6) does not take into account the economic restrictions or costs of the changes in the X_j components. Moreover, $\operatorname{grad} h(\mathbf{x})$ is always orthogonal to the hyperplane (2), and its direction is set by the vector **b**. Due to that, the solution of the problem (3) will be the point lying on the border of the admissible area at the point of its intersection with the vector $\overrightarrow{\mathbf{x}^0\mathbf{x}^1}$, where $\mathbf{x}^1 = \mathbf{x}^0 + a\mathbf{b}$, $a > 0$. If we take into account the economic restrictions and costs of changes in X_j, we will receive the problem

$$\begin{cases} h(\boldsymbol{x}) \to max, \\ x_j = x_j^0 + \Delta_j,\ j = 1,\ \ldots\ ,\ m, \\ \Delta_j \in G_j,\ j = 1,\ \ldots\ ,\ m, \\ v_j(\Delta_j) \le V_j,\ j = 1,\ \ldots\ ,\ m, \end{cases} \tag{7}$$

where $v_j(\Delta_j)$ is the function of costs for changing the j-th component, and V_j is the maximum value of costs for changing the j-th component.

Achievement by the function $h(\mathbf{x})$ of the required probability p_0 with minimum costs for changing the vector of socioeconomic indicators **x**. In this case, the problem can be presented as

$$\begin{cases} \sum_{j=1}^{m} v_j(\Delta_j) \to min, \\ x_j = x_j^0 + \Delta_j,\ j = 1,\ \ldots\ ,\ m, \\ \Delta_j \in G_j,\ j = 1,\ \ldots\ ,\ m, \\ h(\boldsymbol{x}) = p_0. \end{cases} \tag{8}$$

Direct consequences of the set effect on the object are estimated by the value of the target function in the problems (6)–(8). Indirect consequences can be evaluated in each particular case based on the value of the vector of socioeconomic indicators received from the accomplishment of the optimization problem. The main disadvantage of the problems (7) and (8) is the complexity of setting the cost functions $v_j(\Delta_j)$. We can circumvent this restriction by presenting the problem as follows:

$$\begin{cases} \sum_{j=1}^{m} \rho_j \cdot \left(x_j/x_j^0 - 1\right)^2 \to min, \\ h(\boldsymbol{x}) = p_0, \\ x_j, \rho_j \ge 0,\ j = 1,\ \ldots,\ m. \end{cases} \tag{9}$$

The essence of the problem (9) is as follows. It is necessary to achieve the transition of the region to the state $h(\mathbf{x}) = p_0$ with a minimum mean square relative change in the values of the components of the socioeconomic indicators vector. This problem definition does not require explicit specification of the labor intensity of the changes in the X_j indicators. The weighting coefficients ρ_j allow us to take into account the peculiarities of changing the variables, e.g. labor intensity of the change, the importance of the change in each of the variables. If the a priori information about the X_j indicators included in (2) is missing, we take all weighting coefficients ρ_j equal to 1. The zero value of the weighting coefficient can be used where the index of the b_j coefficient has an opposite direction as compared to the expected one (e.g. in this problem, it is the coefficient $\boldsymbol{b_1}$=-2,381<0 with the aggregate birth rate coefficient), which makes it possible to record the actual value of X_j. It should be noted that this can be attributed to the following dependency: the birth rate is higher in regions with a higher standard of living as compared to regions with a low standard of living. This issue cannot be solved within the framework of the optimization problems offered in this publication, as the coefficient vector is determined based on the training sample and reflects the current socioeconomic environment in the country.

We can introduce additional restrictions on maximum changes of the components in the problem (9). In such a case, the optimization problem will look as follows:

$$\begin{cases} \sum_{j=1}^{m} \rho_j \cdot (x_j/x_j^0 - 1)^2 \to min, \\ h(\boldsymbol{x}) = p_0, \\ x_j = x_j^0 + \Delta_j,\ \Delta_j \in G_j,\ j = 1,\ \dots\ ,\ m, \\ x_j \geq 0,\ \rho_j \geq 0,\ j = 1,\ \dots\ ,\ m. \end{cases} \tag{10}$$

Additional restrictions in (10) allow us to factor in the specifics of the changes in the socioeconomic indicators X_j. Let us consider the possibilities of using optimization problems by the example of (9), assuming $\rho_j = 1$, $j = 1, 2, \dots, m$. We will solve the problem (9) using the penalty function method [3]. We will test the logistic regression as the management model for three subjects of the Ural Federal District, namely the Kurgan Region, the Sverdlovsk Region and the Chelyabinsk Region. According to their HDI values, each of these regions has a different standard of living (low, high and medium, respectively). The solution of the optimization problem (10) allowed us to calculate the values of socioeconomic indicators, by reaching thereof the optimum is attained, i.e. the best state of the system ensuring the improvement in the standard of living of the population.

The Kurgan Region can be ranked among the regions with a low HDI. The probability of its classification as a region with a high HDI: $h(\mathbf{x^0}) = 0.005$. Table 3 provides the results of management aimed at reaching the threshold level ($h(\mathbf{x}^*) = 0.5$) for the Kurgan Region, which were obtained by solving the problem (9). To compare the estimate values of socioeconomic indicators, Table 3 contains their actual values for 2013 (the year of initial values for calcualtions) and for 2016, as well as the target values of indicators taken from Decrees of the President of the Russian Federation No. 596–600 and No. 606 dated 07 May 2012.

Table 3. Actual, target and estimate (optimal) socioeconomic indicators of the Kurgan Region.

Indicator	Actual values for 2013 ($\mathbf{X}^0$)	Actual values for 2016	Target values	Estimate values ($\mathbf{X}^*$)
X_1	2.12	2.03	1.753	1.55
X_2	0.68	0.69	0.74	0.70
X_3	5.33	5.44	5.69[b]	6.66
X_4	0.38	0.33	–	0.45
X_5	0.19	0.21[a]	–	0.19
X_6	1.99	1.54[a]	2.7	1.60
X_7	2.71	2.58	1.93	2.26
X_8	4.64	5.5	10.0	4.83
X_9	0.92	1.03	1.00	0.84
X_{10}	0.75	0.84	–	0.67
X_{11}	2.60	2.88[a]	–	2.85
X_{12}	0.21	0.24	–	0.20
X_{13}	1.28	1.75	–	1.21

Note (here and throughout Tables 4-5):
[a] for 2015
[b] the value was calculated proceeding from the established target value of the number of mortgage loans granted in the amount of 815,000 per annum throughout Russia with preserving the same population number as in 2012.
Sources: Rosstat; [8].

The values of socioeconomic indicators obtained by solving the optimization problem show the minimum level that would allow a constituent entity of the Russian Federation to increase its HDI and improve the standard of living of its population. Therefore, the estimate (optimal) values of indicators cannot serve as the new proposed benchmarks, but should be taken into consideration when optimizing the costs and specifying the priorities of further development. For the Kurgan Region, such priorities in improving the standard of living of its population should include: (1) Increasing life expectancy (X_2) up to 70 years (in general, this level was achieved in Russia as early as in 2012, but for this region, with its initially low indicator, this is an optimal and realistic target); (2) Developing mortgage lending (X_3), which has been significantly reduced in recent years as a result of macroeconomic instability; (3) Increasing the investment activity in the region (X_4), which would help kick-start the manufacture, attract free labor and improve the paying capacity of the population; (4) Reducing the unemployment rate (X_{10}) to 6.7%.

The Sverdlovsk Region has a relatively high HDI. The solution results of the optimization problem, which provided for the attainment of a higher HDI corresponding to the increase in the $h(\mathbf{x})$ probability of up to 0.95, are presented in Table 4. According to the calculations, an important condition for probable improvements in the standard of living of the population in the Sverdlovsk Region is economic growth, which is impossible without additional investments (X_4). Development of mortgage lending (X_3) will help raise investments in the construction sector, a driver of the region's economic growth.

Table 4. Actual, target and estimate (optimal) socioeconomic indicators of the Sverdlovsk Region.

Indicator	Actual values for 2013 ($\mathbf{X}^0$)	Actual values for 2016	Target values	Estimate values ($\mathbf{X}^*$)
X_1	1.87	1.91	1.753	1.807
X_2	0.698	0.70	0.74	0.701
X_3	6.153	6.22	5.69[b]	6.403
X_4	0.817	0.799	–	0.868
X_5	0.363	0.411[a]	–	0.371
X_6	2.25	1.97[a]	2.7	2.154
X_7	2.239	2.219	1.93	2.196
X_8	5.28	6.1	10.0	5.315
X_9	0.927	0.97	1.00	0.916
X_{10}	5.9	6.2	–	0.583
X_{11}	3.66	4.24[a]	–	3.729
X_{12}	0.312	0.278	–	0.308
X_{13}	1.696	1.916	–	1.677

The Chelyabinsk Region is classified as a region with a high HDI with a probability $h(\mathbf{x}^0) = 0.805$, which is average for Russian regions, but lower than in the Sverdlovsk region. When solving the optimization problem for improving the standard of living of its population, we considered three options of increasing the probability $h(\mathbf{x})$. The results of solving the problem for three probabilities $h(\mathbf{x}^1) = 0.805$, $h(\mathbf{x}^2) = 0.9$, and $h(\mathbf{x}^3) = 0.95$ are provided in Table 5.

Table 5. Actual, target and estimate (optimal) socioeconomic indicators of the Chelyabinsk Region.

Indicator	Actual values for 2013 ($\mathbf{X}^0$)	Actual values for 2016	Target values	Estimate values ($\mathbf{X}^1$)	Estimate values ($\mathbf{X}^2$)	Estimate values ($\mathbf{X}^3$)
X_1	1.8	1.84	1.753	1.78	1.76	1.72
X_2	0.695	0.705	0.74	0.70	0.70	0.70
X_3	8.16	7.73	5.69[b]	8.28	8.46	8.75
X_4	0.62	0.55	–	0.62	0.64	0.65
X_5	0.25	0.33[a]	–	0.25	0.25	0.25
X_6	2.44	1.86[a]	2.7	2.42	2.39	2.33
X_7	2.32	2.39	1.93	2.31	2.29	2.26
X_8	4.68	5.5	10.0	4.69	4.7	4.72
X_9	0.78	0.90	1.00	0.78	0.78	0.77
X_{10}	0.6	0.71	–	0.6	0.59	0.59
X_{11}	3.67	3.94[a]	–	3.69	3.72	3.76
X_{12}	0.25	0.19	–	0.25	0.25	0.25
X_{13}	1.30	1.76	–	1.29	1.29	1.28

Depending on the region's capabilities and in accordance with Table 5, we can consider various development scenarios of the Chelyabinsk Region. The more favorable dynamics the indicators demonstrate, the higher is the probability of improvements in the standard of living in the region. However, for some indicators (X_2, X_5, X_{10}, X_{12} and X_{13}), there is little or no difference in their values in all the three scenarios, which means that these values are optimal for the regional system analyzed. For the Chelyabinsk Region, mortgage lending (X_3) must become the the priority in improving the standard of living of its population.

5. Conclusion

The values of socioeconomic indicators for Russian regions obtained from solving the optimization problems show that the achievement of individual target indicators of the regional policy alone will ensure the stable regional development and improve the population standard of living. Thus, to improve the standard of living, it is not enough to minimize the mortality rate, but it is crucial to radically change the lifestyle of the population in order to increase life expectancy. As the hazards to health and life of the population are often beyond the scope of direct influence of medicine, but to a large extent depend on nutrition, environment, habits, behaviors, and lifestyles, a higher availability of physicians will not substantially increase the probability of improvements in the standard of living of the population. Significant efforts intended to support reproduction of population will not increase the probability of improvements in people's standard of living. With the current levels of financing in the healthcare and utilities sectors, thestandard of living of the population can be enhanced by improving their efficiency. Reduction in high unemployment rates alone will increase the probability of improvements in the standard of living of the population. Promotion of mortgage lending is crucial not only for economic growth, but also for human potential development. The identified positive relationship between the number of mortgage loans and the HDIs is typical both for the constituent entities of the Russian Federation with a low standard of living and for those with a relatively high one. Investment activities also have a positive effect on the standard of living of the population. However, our calculations demonstrate that not only the volume of investments is important, but also the efficiency of their use. The same GRP increase in a more developed region (with a higher standard of living) requires on average less investments in fixed assets than in a less developed region. Therefore, it is possible to increase the probability of improvements in the standard of living with the existing ratio of investments in fixed assets to GRP.

Acknowledgments

This work has been performed with the support of the Russian Foundation for Basic Research, Grant No. 18-010-00493.

References

[1] Bobylev S N 2012 Indicators of sustainable development for Russia *Bulletin of MGGU. Socioecological technologies* **1** 8–18.
[2] Vorontsov K V 2007 *Lectures on regression recovery algorithms*. URL: http://www.ccas.ru/voron/download/Regression.pdf.
[3] Panteleev A V and Letova T A 2005 *Optimization methods in examples and problems* 2nd edition, revised (Moscow: Vyshaya Shkola) 544 p.
[4] Hilbe J M 2009 *Logistic Regression Models* (US: CRC Press).
[5] Tyrsin A N and Kostin K K 2017 Assessing the logistic regression as an extremum problem *Bulletin of Tomsk State University. Management, computer engineering and computer science* **3** 52–60.
[6] Vasilyev N P and Egorov A A 2011 Experimental calculation of the logistic regression parameters using the Newton – Raphson method to assess the cold resistance of plants *Mathematical biology and bioinformatics* **6(2)** 190–9.

[7] Tyrsin A N, Chistova E V and Kostin K K 2017 Logistic regression as a diagnostic model of stochastic systems *Applied Methods of Statistical Analysis. Nonparametric Methods in Cybernetics and System Analysis*: Proc. of Int. Work-shop (Novosibirsk: NSTU publisher) 355–62.
[8] Grigoriev L M and Bobylev S N 2014 *Report on human development in the Russian Federation for 2014* (Moscow: Analytical Center at the Government of the Russian Federation) 204 p.