

Многокритериальный метрический анализ данных при моделировании человеческого капитала

Г. К. Каменев^{1,a}, И. Г. Каменев^{1,2,b}

¹Вычислительный центр им. А. А. Дородницына Федерального исследовательского центра
«Информатика и управление» Российской академии наук,

Россия, 119333, г. Москва, ул. Вавилова, д. 40

²ФГАОУ ВО Национальный исследовательский университет «Высшая школа экономики»,

Россия, 101000, г. Москва, ул. Мясницкая, д. 20

E-mail: ^a gkk@mail.ru, ^b igekam@gmail.com

Получено 19.12.2019, после доработки — 11.07.2020.

Принято к публикации 17.07.2020.

В статье описывается вычислимая модель человека в информационной экономике и демонстрируется многокритериальный оптимизационный подход к метрическому анализу модельных данных. Традиционный подход к идентификации и исследованию модели предполагает идентификацию модели по временным рядам и прогнозирование дальнейшей динамики ряда. Однако этот подход неприменим к моделям, некоторые важнейшие переменные которых не наблюдаются явно, и известны только некоторые типичные границы или особенности генеральной совокупности. Такая ситуация часто встречается в социальных науках, что делает модели сугубо теоретическими. Чтобы избежать этого, для (неявной) идентификации и изучения таких моделей предлагается использовать метод метрического анализа данных (ММДА), основанный на построении и анализе метрических сетей Колмогорова–Шеннона, аппроксимирующих генеральную совокупность данных модельной генерации в многомерном пространстве социальных характеристик. С помощью этого метода идентифицированы коэффициенты модели и изучены особенности ее фазовых траекторий. Представленная в статье модель рассматривает человека как субъекта, обрабатывающего информацию, включая его информированность и когнитивные способности. Составлены пожизненные индексы человеческого капитала: креативного индивида (обобщающего когнитивные способности) и продуктивного (обобщает объем освоенной человеком информации). Поставлена задача их многокритериальной (двухкритериальной) оптимизации с учетом ожидаемой продолжительности жизни. Такой подход позволяет выявить и экономически обосновать требования к системе образования и социализации (информационному окружению) человека до достижения им взрослого возраста. Показано, что в поставленной оптимизационной задаче возникает Парето-граница, причем ее тип зависит от уровня смертности: при высокой продолжительности жизни доминирует одно решение, в то время как для более низкой продолжительности жизни существуют различные типы Парето-границы. В частности, в случае России применим принцип Парето: значительное увеличение креативного человеческого капитала индивида возможно за счет небольшого снижения продуктивного человеческого капитала (обобщение объема освоенной человеком информации). Показано, что рост продолжительности жизни делает оптимальным компетентностный подход, ориентированный на развитие когнитивных способностей, в то время как при низкой продолжительности жизни предпочтительнее знаниевый подход.

Ключевые слова: многокритериальная оптимизация, метрические сети, визуализация данных, человеческое развитие, идентификация модели, метод достижимых целей, интерактивные карты решений, человеческий капитал, метрический анализ данных

Работа выполнена при частичной поддержке РФФИ, проект 18-01-00465 а «Разработка методов многомерного метрического анализа социальных данных». Авторы также благодарят Е. В. Доброва за помощь в переводе.

UDC: 51-77, 517.977.5, 519.86, 303.094, 330.45, 331.101.26

Multicriterial metric data analysis in human capital modelling

G. K. Kamenev^{1,a}, I. G. Kamenev^{1,2,b}

¹GBI Federal Research Center of Computer Science and Control Dorodnitsyn Computing Centre
of the Russian Academy of Sciences,
40 Vavilova st., Moscow, 119333, Russia

²FGAEI HE National research university “Higher school of economics”
20 Myasnitskaya st., Moscow, 101000, Russia

E-mail: ^a gkk@mail.ru, ^b igekam@gmail.com

Received 19.12.2019, after completion — 11.07.2020.

Accepted for publication 17.07.2020.

The article describes a model of a human in the informational economy and demonstrates the multicriteria optimizational approach to the metric analysis of model-generated data. The traditional approach using the identification and study involves the model's identification by time series and its further prediction. However, this is not possible when some variables are not explicitly observed and only some typical borders or population features are known, which is often the case in the social sciences, making some models pure theoretical. To avoid this problem, we propose a method of metric data analysis (MMDA) for identification and study of such models, based on the construction and analysis of the Kolmogorov–Shannon metric nets of the general population in a multidimensional space of social characteristics. Using this method, the coefficients of the model are identified and the features of its phase trajectories are studied. In this paper, we are describing human according to his role in information processing, considering his awareness and cognitive abilities. We construct two lifetime indices of human capital: creative individual (generalizing cognitive abilities) and productive (generalizing the amount of information mastered by a person) and formulate the problem of their multi-criteria (two-criteria) optimization taking into account life expectancy. This approach allows us to identify and economically justify the new requirements for the education system and the information environment of human existence. It is shown that the Pareto-frontier exists in the optimization problem, and its type depends on the mortality rates: at high life expectancy there is one dominant solution, while for lower life expectancy there are different types of Pareto-frontier. In particular, the Pareto-principle applies to Russia: a significant increase in the creative human capital of an individual (summarizing his cognitive abilities) is possible due to a small decrease in the creative human capital (summarizing awareness). It is shown that the increase in life expectancy makes competence approach (focused on the development of cognitive abilities) being optimal, while for low life expectancy the knowledge approach is preferable.

Keywords: multicriteria optimization, metric net, data visualization, human development, model identification, feasible goals method, interactive decision maps, human capital, metric data analysis

Citation: *Computer Research and Modeling*, 2020, vol. 12, no. 5, pp. 1223–1245 (Russian).

The reported study was co-funded by RFBR, project number: 18-01-00465 a “Development of social data multidimensional metric analysis methods”. The authors are also grateful to E. V. Dobrov for the help with translation.

Introduction

The article describes the dynamic model of the human capital's evolution, identified and researched using the Method of metric data analysis (MMDA). The variables used in the model are not observable in social data. Therefore, both identification and research of the model are non-trivial. For these purposes, we use a special method of analysis based on model experiments.

A number of scientists researched human capital (the classic ones are [Mincer, 1958; Becker, 1964; Schultz, 1971; Bourdieu, 1986; Coleman, 1988; Mincer, 1989]). The main scientific ideas of these researches are: the interpretation of human capital as an asset that creates income for the owner, and the corresponding investment modeling, and the usage of lifelong indicators of return on (human) capital. This article uses this classical fundamental features of human capital and, if possible, classical terminology and formalism.

Modern researches on human capital [Galor, Weil, 2000; Galor et al., 2017; Currie, Almond, 2011; Crook et al., 2011; Ployhart, Moliterno, 2011; Fleisher et al., 2010; Davies et al., 2018] usually focus on the problem of parental household' choice or apply human capital's index-based macroeconomic operationalization, without giving priority to its age dynamics (its increase and especially decrease). Modern Russian University science [Radaev, 2002; Gimpelson, 2016; Gimpelson et al., 2007, 2015, 2016; Zabelina et al., 2014; Shabunova, Leonidova, 2011; Kapelyushnikov, 2008, 2010, 2011, 2012; Smirnov et al., 2005; Geliga et al., 2011; Koritsky, Grigorova, 2010] generally continues the international paradigm of human capital.

The approach to the study of the human capital of the population (country) based on the choice of the parent household includes a set of explicit and implicit assumptions that were valid in the second half of the XX century, but are now outdated and make it limitly applicable to modern problems of human capital management:

1. It was assumed that the speed of learning work experience does not depend on the level of education. This is true if we consider only the classical knowledge-based approach to education, but as the modern competence approach to education spreads, this interpretation becomes outdated.
2. It was assumed that people receive education mostly before the first employment (being funded by parents and/or government). This concept corresponded to the realities of the second half of the XX century, but today the concept of continuous learning (lifelong learning, regular professional development, etc.) is generally accepted. A significant contribution to the funding of continuing education is made by workers and their employers.
3. It was assumed that human capital increases with age and experience. This assumption was true for the moderate life expectancy and work seniority of the XX century. However, in modern post-industrial economies, where the duration of employment is not up to 20, but regularly more than 40 years, it is increasingly important to recognize qualification reduction in older age. This effect cannot be described by a linear or logarithmic function.

These problems lead to the fact that the micro-basis of human capital from models with monotonous dynamics becomes ineffective for justification of the tasks of the firm and the government. Instead, index-type models of human capital are actively used, while fundamentally lacking micro-basis (see, for example, [Fleisher et al., 2010]), but are better suited to justify macroeconomic decisions. Many Russian authors ([Zabelina et al., 2014; Shabunova, Leonidova, 2011] etc.) also actively calculate the index estimate of human capital and use it in interregional/cross-country comparisons of human capital.

In index models formed aggregate indicator of human capital that characterize the capital (the ability of a person to create a useful product), and the industries traditionally associated with human

capital creation (education, health, etc.)). The disadvantage of this approach is associated with the vulnerability of any of the aggregated formulas to the essential criticism (in this formula values with different dimensions and dimensionless values are summarized, and the weighting coefficients have no meaningful interpretation).

An alternative to this approach is to construct functions of human capital that demonstrate nonlinear non-monotonic dependence on various dynamics factors. We offer our own approach to quantitative analysis of human capital, which allows us to link the characteristics of the population (country) with the characteristics of specific individuals: their cognitive abilities and awareness. We suggest using simulation mathematical modeling and multi-criteria metric analysis of model-generated data. Our approach to the study of human capital has no direct analogues, since previously such models had two specific problems in the field of mathematical modeling: low quality of identification and lack of analytical solutions. Let's look at these problems thoroughly.

One of the main directions of multicriteria analysis application is the problem of optimal (rational) choice/decision made by people. In this kind of problems, there is often a dynamic model describing the consequences of a choice made. A specific feature of the choice problems made by people is the existence of implicitly given variables (characteristics) that are not directly observed for the most of time. This leads to the fact that the problem of optimal choice in the Social Sciences is divided into sociometric (econometric), in which the model includes mainly clearly observed variables that don't fully reflect the substantive interrelations, and theoretical models where the variables and the choice logic justified substantially, but are not comparable with factual data.

The proposed method allows us to research multicriteria optimal choice based on model experiments while comparing it with individual (particular) restrictions on unobservable variables known from social statistics.

Models with such features are most common in applied mathematical fields, considering rational human choice with long-term consequences. One of the special cases chosen for this research is the modeling of the choice for education. Here there is a strategic choice, the delayed result (school and higher education affects the entire future life of a person, and any later changes in the life strategy are mostly difficult) and requires prioritization due to the internal incompatibility of some goals. At the same time, the essential characteristics of a person (the amount of information he has mastered and his cognitive abilities) are not observable and therefore cannot be compared with a statistical data series.

Further, while studying any model of choice with nonlinear nonmonotonic dynamics, the main difficulty is that it's systems of equations or inequalities may be analytically unsolvable. Firstly, in such dynamic system, intercorrelation effects occur (variables-functions being simultaneously variables-factors, for example: more developed cognitive abilities increase the speed of learning, and knowledge of certain thinking techniques develops cognitive abilities). Secondly, at different stages of life some factors of dynamics can change the nature of their impact, even to the opposite (e.g., knowledge helps the individual to build the systemic understanding of the world faster in his youth, but in his old age, the presence of such an understanding makes it difficult to acquire new knowledge not compatible with this whole picture).

Thus such models with non-monotonic dynamics are computable, but it is difficult to obtain their analytical solution (a direct expression of all phase variables for any given initial values along sight the entire trajectory). Therefore, it is difficult to prove the existence of a single (global) optimum for the choice problem (especially if it is multicriterial). In the considered model of human capital, this leads to the appearance of the Pareto-frontier. Mathematical methods for research of this type of models have not previously been applied to human capital models, so our study is of interest not only as a new class of human capital models, but also as an example of using simulation methods (model experiments, construction and approximation of metric nets, identification and visual analysis of the Pareto-frontier and of the set of Pareto-efficient solutions), in a new subject area.

There are no time series from social statistics for model variables since they are not observable. However, there are a number of known restrictions that they must meet. We use a special method that allows us to construct (approximate) the set of all valid combinations of model parameters for which the trajectories of phase variables satisfy the corresponding constraints. It turns out that this set of acceptable parameters is narrow enough to solve the problem of model calibration. Note that the fact of good localization of model parameters through implicit restrictions on the object state is also evident in other areas of science (it is enough to mention the so-called anthropic principle in cosmology [Barrow, Tipler, 1986] or, on a smaller scale, numerous physical experiments aimed to identify the range of possible values of fundamental constants, subatomic particles' characteristics, etc.). In our case, the reliability of model calibration is determined by the accuracy of approximation of the set of acceptable combinations of parameter values.

This issue will be briefly discussed in the section "Model identification and calibration with socio-psychological population requirements". Model experiments allows us to obtain the values of such variables over the entire trajectory, and they can be compared with individual boundary values for different ages, as well as with patterns of development of the entire population as a whole known from qualitative sociological research. The proposed method allows us to study the multicriteria optimal choice based on model experiments while comparing it with individual (particular) restrictions on unobservable variables known from social statistics.

This approach (including the method of analyzing metric data developed by us), being currently actively developed (mainly in populational and environmental modeling, see, for example [Kamenev et al., 2016], as well as in classical economic modelling [Kamenev, Olenev, 2015]), allows us to combine various effects known from laboratory psychological research into a new class of models of human capital.

Let us emphasize, however, that in our study we consider the model of human capital of a population (country) with a micro-basis at the level of the person as a whole, and not the model of the human psyche. We do not aim to describe completely all the mental processes. We only combine known psychological effects that are significant from the point of view of the nonlinear age dynamics of human capital. That is why mental processes directly related to cognitive activity are presented in our model in an aggregated form (as the development and usage of information), and other processes are not considered. This higher level of generalization is due to the subject of research and distinguishes the proposed model from numerous quantitative models in psychology that study the features of individual mental processes (attention, memory, perception, etc.).

The methodological basis of the multi-criteria metric analysis of social data

Let us briefly consider the general logic of the research method used in this article. The dynamic model is given in discrete time, assuming natural development of several (two in this research) characteristics of the studied developing object (person). The value of each characteristic can be calculated from the known equations of the model for any period of time at known starting values. Starting values, in turn, are considered as controls, since their values are the result of the choice of the subject. In this research, we initially describe the developing object as if it is simultaneously the subject of decision-making. I.e. the person (family, household) independently makes a multicriterial optimizational choice about his starting characteristics based on his interests. In other words, it is assumed that the choice is made in the interests of the developing person. We will introduce a different formulation in another publication, in which the decision is made by the government, focusing not on the individual interests, but on the entire population.

The functionals given on the characteristics are choice criteria. In particular, to typical functionals can be assigned a maximum of features and the integral over the time of life (both types are described

in the article). The method allows setting any formulation with limited internal resources, additionally bounding the space of choice.

The primary element of the study is a model experiment time series (observation), the totality of which (the model generated data) forms model-generated data set. It is used to characterize the general population of possible states of the studied object in time with given completeness and reliability.

It is assumed that the phase variables used in the model cannot be compared with statistical series, as they are not directly observable, but some specific quantitative characteristics of variables and/or functionals obtained from qualitative social studies are known. Their presence allows for a procedure substantially similar to the identification (calibration) of the model.

The approach applied in this study (Method of Metric Data Analysis, MMDA [Kamenev, Kamenev, 2017, 2018]) allows us to identify a mathematical model using some separated social data and study its features (including the Pareto-boundary). The mathematical basis of the method consists in constructing (ε, δ) -nets and (ε, δ) -coverings, introduced by Claude Shannon [Shannon, 1948] for the general population of humans as information carriers, providing topological and metric analysis of its structure and its usage in criteria space.

A universal method for “coding” or “tabulating” nonconvex sets is the construction of metric ε -nets [Kolmogorov, Tikhomirov, 1959]. Geometrically, this means that a set is approximated by a collection of simple figures (such as balls or cubes of appropriate dimension), whose diameters are taken smaller to achieve a higher approximation accuracy. A metric ε -net of a set is its subset such that any point of a set lies at a distance of at most ε from the subset. If “balls” are constructed in the considered spatial metric around the points of a metric ε -net, then they cover the entire approximated set forming ε -covering. This corresponds to the set approximated by a collection of usual balls in the Euclidean metric with the norm $\|z\| = (\sum z_j^2)^{1/2}$ and by a collection of cubes in the Chebyshev metric with the norm $\|z\| = \max\{|z_j|\}$ (in this research); or any other metrics being relevant to the object. For each point of such a collection, we can quickly find a “true” point of the set (the nearest point of the metric ε -net) separated from it by a distance of at most ε in the considered metric. Moreover, for smaller ε , the true points are closer to the points of the approximated set and there are fewer “redundant” points in the approximating collection of balls, but a greater number of points have to be constructed in the net. The optimal complexity of ε -nets constructing, is characterized using the concepts of metric ε -entropy.

Bounded implicitly defined sets (images of a maps) can be approximated by applying adaptive stochastic techniques based on the Deep Holes method (DHM) [Kamenev, 2001]. Specifically, for a set defined implicitly by a nonlinear “black-box” mapping, these techniques are used to construct a covering that approximates it with prescribed accuracy ε and completeness η . The completeness of the approximation is characterized by the fact that only a small part (a part with a small measure) of the approximated set lies outside the constructed collection of balls with centers at points of the net. To estimate the completeness η of the ε -covering (the measure of the covered set) we use a test sample consisting of M points from the pre-image set and check the sampling fraction η^* that belong to the tested covering. The collection of balls with accuracy ε and completeness η forms so-called (ε, δ) -coverings, $\delta = 1 - \eta$ [Shannon, 1948]. In this approach the reliability of (ε, δ) -covering depends on the value of M (the probability of proximity of η and η^*) [Kamenev, 2001].

The stochastic implementation of the DHM allows us to construct (ε, δ) -coverings with given precision (ε), completeness (δ) and reliability characteristics. The data for MMDA and DHM implementation was generated using the calibrated model. We visualize the General population’ (ε, δ) -coverings by Interactive Decision Maps (IDM) [Lotov et al., 2004]: 2D and 3D maps and atlases of multidimensional explicitly-given sets. This maps are also used for solving optimizational problem in Feasible Goals Method (FGM) [Lotov et al., 2004]: visual analysis of Pareto-frontier and construction of corresponding optimal solution. For identification purposes, we used the version of MMDA called

the Identification Sets Method (ISM, see [Kamenev, 1998, 2016]) to approximate the calibration parameters set (called in the ISM method as “identification set”), for which the trajectory tube satisfies the known calibration requirements. In this approach the reliability of (ε, δ) -covering depends on the value of M (the probability of proximity of η and η^*) [Kamenev, 2001]. More precisely, let us denote probability defined on the space of samples of the volume M as P . Then the reliability $P\{\eta > \eta^* - \Delta\}$ of $\eta^* - \Delta$ to be lower estimate of the completeness η of a given coverage is evaluated by the following result:

Theorem 1. $P\{\eta > \eta^* - \Delta\} \chi(\eta, M), \chi(\eta, M) = 1 - \exp(-2M\Delta^2)$.

This theorem [Kamenev, 2001] characterizes the sample size M needed to construct an approximation with a given completeness η and reliability χ .

This article contains an enhanced and extended description of the study, which first results are published in [Kamenev, 2018; Kamenev, Kamenev, 2018]. Unlike previous publications, it not only presents the model and the task of individual, but also describes the most important model’s features, including the individuals’ human capital’s assessment with different starting parameters (controls).

Model introduction, dependencies and dimensions analysis

We denote a person’s age as $n \in N$, where $N \subset \mathbf{N}$ is a set of allowable values of the person’s age. The time dynamics in the model is considered to be identical to the age dynamics and therefore is discrete. All of the model’s constructs and transformations can be transferred in continuous time, but it is not relevant due to the discrete nature of the vast majority of social data and their accuracy level. For example, the age of a person in every social data is specified with the accuracy of one year with rounding down (i.e. the person, whose age tends to n from below, is represented in the social statistics data as having the age of $n - 1$). It cannot be neglected, because the age dynamics refers to the main subject of the study.

Next, we denote a person’s cognitive abilities as $x \in X$, $X \subset \mathbf{R}$, is a set of allowable values of the cognitive abilities x of different people. Positive values of x mean that the human mentality develops, while negative values mean that it degrades. The subset of values, taken in a certain age $n + 20$, is denoted as X_n .

Further, $y \in Y$ – accumulated information which can be applied by a person, consciously or unconsciously (is available in memory). Formalizations for $Y \subset \mathbf{R}$ are similar to those, which were mentioned above, for X .

Let’s consider the nature of the mutual dependency of cognitive abilities and human awareness. On the one hand, the human mentality is a relatively stable system, the characteristics of which are accumulated in time. Thus, the basis for any model of awareness and cognitive abilities should be their transfer to the next period of time and, therefore, to the next age. On the other hand, awareness and cognitive abilities affect each other in different ways.

- $x_{n+1} = x_n + f_1$, where f_1 are cognitive abilities change factors;
- $y_{n+1} = y_n + f_2$, where f_2 are awareness change factors.

Let’s take a closer look at the factors of changes in cognitive abilities and awareness used in the model, including their psychological basis.

The very definition of cognitive abilities indicates that they increase the effectiveness of the new information obtaining. Therefore, we introduce cognitive abilities into the model as a factor of awareness change. On the contrary, the new information obtaining depends on the amount already obtained (through analogies, associations, etc.): for erudite personality it is much easier to remember new facts and he has a large number of relationships between them. This factor’s significance is

confirmed by a large number of psychological studies (for the most complete overview of the factor, see, for example, [Anderson, Bower, 2014]).

To take into account the mutual influence of the considered abilities, we will change our approach from the traditional linear dependencies to the bilinear approximation, which is justified on such small increments as one year according to the Taylor series of factors f_1 and f_2 . Then the full function of the age dynamics of awareness takes the form: $y_{n+1} = y_n + by_nx_n$, where $b > 0$ is the calibration constant.

Factors of cognitive change require a more complex interpretation. First of all, cognitive abilities can be used by the personality for reflection: rethinking of its activities, new cognitive schemes mastering. Reflection, i.e. the usage of cognitive abilities to re-evaluate person's knowledge, abilities, and actions, allows the individual to develop his cognitive abilities. The interrelation between reflection and cognitive abilities is described in a variety of psychological researches, among which the most well-known (and most actively criticized) is the CRT-test model [Frederick, 2005]. It is priority to point for the purpose of our research, that this dependence is nonlinear, since reflection accelerates itself. This can be modeled by a quadratic dependency.

However, the assumption of the continuous and self-accelerating nature of cognitive development would come into fundamental conflict with the available factual social data, which show that by the end of a person's life his cognitive abilities are rapidly declining. In our model, we proceed from the assumption that this leads to the volume of human assimilation of information. A person has to spend mental resources on maintaining connections in memory, i.e. to prevent forgetting. This effectively reduces the cognitive abilities available to a person. The forgetting effect and the forgetting curve are also described in detail in the psychological literature (see, for example, [Averell, Heathcote, 2011; Henderson, 1999]). In bilinear approximation we conclude that the full function of the age dynamics of cognitive abilities takes the form: $x_{n+1} = x_n + a_1x_n^2 - a_2y_n^2$, where $a_1 > 0$ and $a_2 > 0$ are calibration constants.

Thus, our model includes aggregation of mental processes into two variables. Their dynamics is determined by four well-known in the psychological literature effects. Obviously, the model could be expanded with a large number of effects, but it will be shown later (while model identification) that these four effects are sufficient: they allow us to reproduce the most significant (for the purposes of this research) patterns in the development of the human psyche. In turn, the strength of the effects (the values of calibration constants) cannot be quoted from any psychological research, because we use a higher level of aggregation of the psyche. Accordingly, the calibration constants are found by us independently during the model identification.

Taking the age of a person changing in proportion to the absolute time, we obtain two dynamic equations:

$$x_{n+1} = a_1x_n^2 - a_2y_n^2 + a_3x_n, \quad (1)$$

$$y_{n+1} = b_1x_ny_n + b_2y_n. \quad (2)$$

Here, $n = 0, 1, \dots, 100$ are years of a person's life. The validity of the model is based on assumption that person has completed primary socialization and has adult psyche (psyche development during primary socialization is mostly determined by external factors, which cannot be researched with our model's architecture). In order to completely exclude the child or adolescent psyche from consideration, $n = 0$ means the age of 20 years.

The positive values of x mean that the human psyche is developing, while the negative values mean that it is degrading. Cognitive abilities' value is defined in the range from 1 to -1 . The model is calibrated in a way that the population cognitive abilities maximum is close to 1 (it is assumed that person cognitive abilities have a physiological limit that was equated to 1). Thus, the indicator $x \in X$ is relative and dimensionless. Cognitive abilities boundary -1 is interpreted as an intellectual dysfunction (person's intellectual ability get lost, for example, in coma or death). Accordingly, the boundary from

above is rigid (i.e., cannot be exceeded in the model), the boundary from below is soft (its crossing closes the modeled period of person's life).

The indicator $y \in Y$ is not bounded above, but its minimum is 0. The bottom line is "soft": the achievement of 0 is interpreted as a memory dysfunction and, accordingly, brain death. The indicator $y \in Y$ is non-relative (absolute) and, therefore, has a dimension (units of information). At the same time, the y specific values are not important, since a linear substitution can always be made to change the scale.

Due to the dimensionless nature of X , all coefficients are also dimensionless except one (a_2 has dimension $1/\text{information_unit}^2$).

The quadratic dependence displays the self-accelerating nature of reflection: if a person has the opportunity to use psychic resources for self-development, then it occurs quickly due to the multiplicity of neural connections; if he cannot allocate enough resources to maintain the psyche, then it will lead to accelerated degradation (also due to the multiplicity of connections: any element malfunction disturbs the others' work).

The parameters a_3 and b_2 are taken as 1 because they reflect the dynamic nature of the system (the human psyche). It is assumed that the psyche of the next period is the psyche of the previous period, which some endogenous changes. Would we include biological aging as a mental development model factor, then a_3 and b_2 should become less than 1, but more than 0, obviously, closer to 1. However, the current neurophysiology state does not allow us to unambiguously associate aging with mental processes and does not give us social data that would unambiguously indicate the necessity for incomplete psyche transfer to the next period.

All the model's coefficients are assumed to be constant, their values should be determined in the model study on the stage of calibration.

Model controls and optimization criteria

After the calibration of the model it is possible to introduce two possible formulations of the optimization problem:

1. Analysis of the particular individual's human capital dynamics trajectory.

In this problem we are interested in various features of trajectories, but, first of all, the functionals given as $x_{\max} = \max\{x_n\}$ and $y_{\max} = \max\{y_n\}$ on the trajectories of x_n and y_n , creating multicriteria optimization problem $(x_{\max}, y_{\max}) \rightarrow \max$. Basic controls are $x_0 \in X_0$ and $y_0 \in Y_0$, i.e. in the sets of possible starting values of x and y ($n = 0$, i.e., the age of 20 years). Further development is completely computationally explicit, i.e. $x_n = f_3(x_0, y_0, n)$ and $y_n = f_4(x_0, y_0, n)$. Given the model parameters, these dependencies implicitly describe the tube of trajectories in three-dimensional space (n, x, y) from $n = 0$ with all possible for the population variables x_0 and y_0 combinations. This variables and functionals can be calculated, which allows us to visualise their Pareto-frontier using MMDA. This formulation is basic because it reflect the patterns of age dynamics.

2. Analysis of the individual's human capital, taking into account the probability of mortality.

The subject of the first formulation is the dynamics of qualification (cognitive abilities and awareness), which allows to study its key patterns. However, in order to solve the optimization problem (to determine the optimal starting characteristics for the individual), it is necessary to take into account that not everyone lives up to the maximum age (120 years in the model). Since health as a characteristic of the human capital does not belong to our subject of study, the probability of survival to the particular age will be exogenous (seria) variable $p(n)$.

It allows us to use the lifelong indicators of human capital (which study is strongly recommended by the classics of the human capital theory [Becker, 1964; Schultz, 1971]). This problem will be discussed in the more detailed way after the calibration of the model.

Model identification and calibration with socio-psychological population requirements

Our model contains a number of coefficients (constants), which values must be chosen in such a way that the model dynamics correspond to real social data. The variables interdependence greatly complicates the identification of coefficients for which the dynamics of variables (x_t and y_t) reflecting real human development patterns known from (social) psychology.

To establish these limitations, we have studied numerous psychological publications examining human cognitive abilities in its dynamics. This area of psychology is highly debatable: there are many contradictory works. For the purposes of this research, we rely upon the following assumption: person's cognitive abilities degrade with age, but their peak is reached not as fast, as shown by early psychology studies, but at a mature age. If necessary, the model and our research method allow us to recalibrate the model in accordance with any new cognitive abilities concept.

The psychologists and sociologists work on the cognitive abilities dynamics and measurement, for example, IQ tests, is extremely controversial [Salthouse, 2013]. In general, newest researches moves the cognitive abilities peak significantly ahead into the later ages. Therefore, we take $X_0 = [x_0^{\min}, x_0^{\max}]$, where $x_0^{\min} = 0.44$, $x_0^{\max} = 0.80$ (all x_0 must belong to the specified range), and $\operatorname{argmax}\{x_{20}: x_{20} \in X_{20}\} \approx 1$, i.e. cognitive abilities population maximum is achieved by most intellectually developed individuals at $n = 20$, after which abilities' decline begins (we recall that in our scale it is 40 years). Also, we recall the fundamental restriction $x_n \leq 1$ with $\max\{x_n: n = 0, \dots, 100, x_0 \in X_0\} \approx 1$ established by relativity of X .

The total starting amount of information mastered by a person y_0 is a hardly verifiable characteristic and additionally, it is easily scaleable. It is accepted that in the starting (20 years' age old) population it should differ by 25–50%. Preliminary calibration research (in this paper we are skipping it) showed that for the considered problem it is convenient to establish its scale through the range $Y_0 = [y_0^{\min}, y_0^{\max}]$, where $y_0^{\min} = 0.23$, $y_0^{\max} = 0.33$. The awareness dynamics is not specifically limited in this model.

The selection of a_1 , a_2 , and b_1 ensures that the model achieves realistic values. In addition to the cognitive abilities peak, realism criteria also include:

- negative value in X (cognitive abilities) of 80-year-olds ($t = 60$);
- positive value in X (cognitive abilities) of 40-year-olds ($t = 20$).

Such restrictions ensure that the cognitive abilities dynamics correspond to patterns observed in real society (physiological anomalies such as cognitive mental retardation are not taking into account).

Model identification includes the selection of such values of calibration coefficients a_1 , a_2 and b_1 (reflecting the strength of mental development factors in the model), with which the model's dynamics is realistic. Let us describe the scheme of calibration process in MMDA. It uses ε -covering construction for the identification set — the allowable calibration coefficients values combinations. The coefficients from this set creates the trajectories tube satisfying given conditions for $S = X_0 \times Y_0$ starting values.

As mentioned above, the identification set is approximated using stochastic DHM, based on global random probing of the set of all possible model coefficients values combinations for which all trajectories from S meet the required calibration conditions. For a given model coefficients values combinations we check this requirement with the following procedure.

We define a template on the set $S = \{0.44 \leq x_0 \leq 0.8, 0.23 \leq y_0 \leq 0.33\}$ as a uniform square net of 11 nodes for each of the variables. For each combination of coefficients from the ε -coverage of the identification set, a complete tube of trajectories for all ages is constructed, including trajectories for all ages for each of the $\|X\| = 121$ nodes of the net, approximating different initial values of cognitive abilities and awareness. If at least one of the constructed trajectories do not match the calibration criteria, then this combination of calibration coefficients is unrealistic.

A strict mathematical description of this procedure is presented in the publication [Kamenev, 2016], and its application to our model is presented in the publication [Kamenev, 2018]. The approximation of the sets presented in this paper was based on a uniform probability measure defined in the space of parameters and the starting points of the trajectories, with $M = 5000$, which corresponds $\chi = 0.98$ and $\eta = 0.98$. This means that the effects with a probability measure not exceeding 2% remained beyond analysis and characterizes the reliability of the model calibration method.

In Fig. 1 (ε, δ)-covering of the parameters identification set is visualized for which a tube of $11^2 = 121$ starting at the template nodes trajectories satisfies the specified calibration requirements. Note that the set of allowable values of parameters is non-empty, but narrow: the spread of acceptable values for each coefficient is less than 5% (or less taking into account their mutual limitations. This means that the applied calibration criteria are very stringent. If the set would be empty, this would indicate that the model is not suitable for reality describing: in other words, it cannot reproduce the observed dynamics without additional factors inclusion. If the identification set would be significantly larger (a large parameter spread would be allowed), it would indicate a weak predictive power of the model, since the forecast would change significantly with other acceptable parameters values. We choose some point of the ε -net – a combination of parameters $\lambda^* = (0.0210, 0.0650, 0.0415)$, which ensures the calibration conditions fulfillment.

Now it is necessary to check the calibration requirements on the entire trajectory tube (see Fig. 2 on the right), not just on the template. At this stage, the Deep Hole Method [Kamenev, 1998, 2018] constructs a trajectories tube (phase tube) for the entire given initial set and we check the given constraints (see Fig. 2 on the left). The result proves calibration quality. Thus, as a result of identification, we determine the values of calibration coefficients (reflecting the strength of the factors included in the model), measured with more than 95% accuracy and a reliability of 95%. This accuracy measurement is sufficient for the model to be used further in the research of human capital's population dynamics, which we directly describe by human awareness and cognitive abilities.

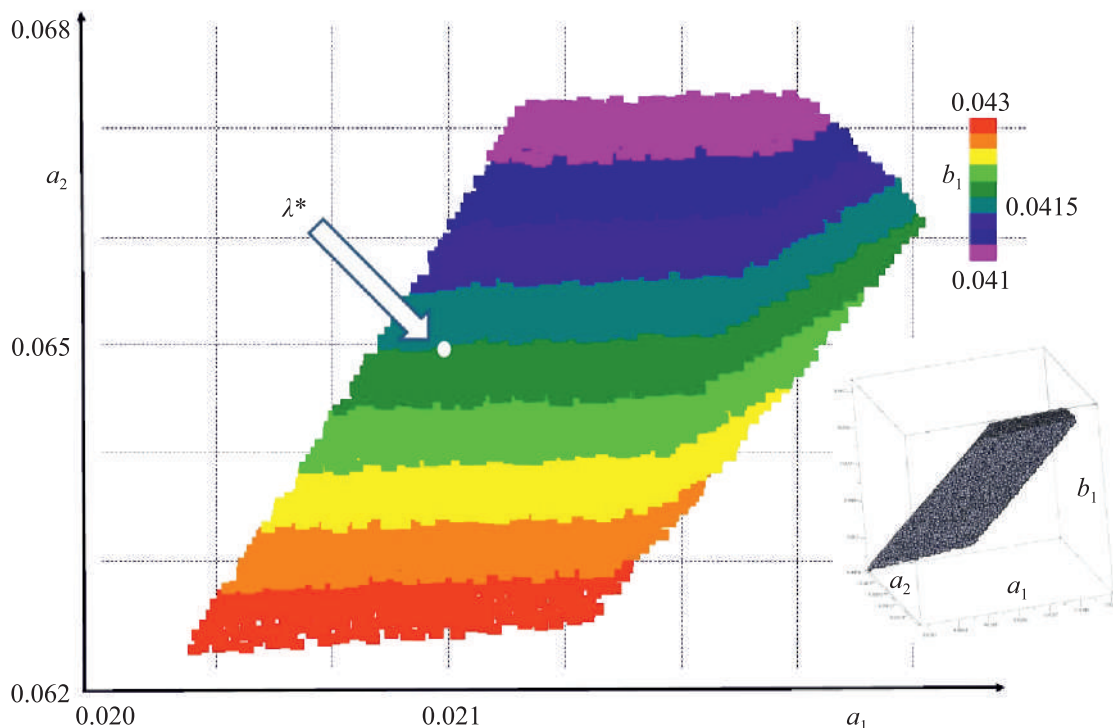


Figure 1. The identification set (set of allowable values) for the 3 calibration parameters combinations (a_1, a_2, b_1)

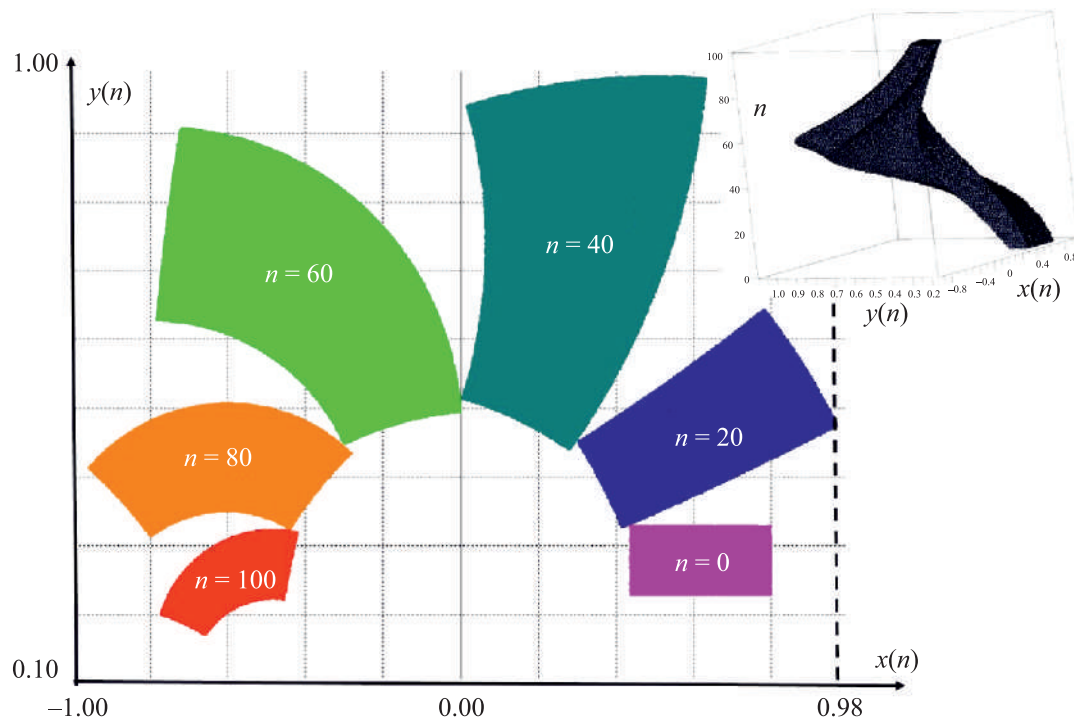


Figure 2. The phase tube of trajectories for cognitive abilities x_0 and awareness y_0 from S , age n based on calibrated parameters λ^* (on the right) with slices of the tube for ages conditions (on the left)

Main features of the phase trajectories

Let us define trajectory x_n as $x(x_0, y_0, n)$, $n = 0, 1, \dots, 100$. Accordingly for y_n , $y(x_0, y_0, n)$. Thus, general population G is a trajectory tube:

$$G = \{(x, y, n) \in X \times Y \times N : x = x(x_0, y_0, n), y = y(x_0, y_0, n), n = 0, 1, \dots, 100, (x_0, y_0) \in S\}. \quad (3)$$

The calibrated model has been researched in the context of the above-mentioned problem statements. Mathematically, the study also relies on the MMDA and, therefore, involves the construction of exploratory (ε, δ) -nets and (ε, δ) -coverings for the G in the space (X, Y, N) of model-generated data with high metric accuracy ε and completeness $\eta = (1 - \delta)$, which provides topological and metric analysis of the structure of G (see Fig. 2).

The independent interest is the ratio of different states of mentality and the following forecasts about the consequences of different initial values of cognitive abilities and awareness for the autonomous dynamics of the individual as the carrier of information. The visualization of the results is possible on the basis of the interrelated indicators: the values of x and y indicators in different years of life.

Since the calibrated model is dynamic, its capabilities can be identified using trajectory analysis. To do this, it is necessary to construct a multidimensional space of indicators characterizing trajectories. Let us define A as a trajectory given by the combination of values of $x_0 = x_0^*$ and $y_0 = y_0^*$, where $x_0^* = x_0^{\max}$ and $y_0^* = y_0^{\min}$. The main dynamic patterns of both x_n and y_n are relatively simple: they increase until they reach some (single) maximum value, after which they begin to decrease (see Fig. 3). This features can be observed for any x_0 and y_0 combinations.

In the analysis of trajectories, the indicators specified as a functionals on the trajectory are of interest: $x_{\max} = x_{\max}(x_0, y_0) = \max\{x_n\} = \max\{x(x_0, y_0, n)\}$ and $y_{\max} = y_{\max}(x_0, y_0) = \max\{y_n\} = \max\{y(x_0, y_0, n)\}$. To confirm the identification results, indicators corresponding to different ages are

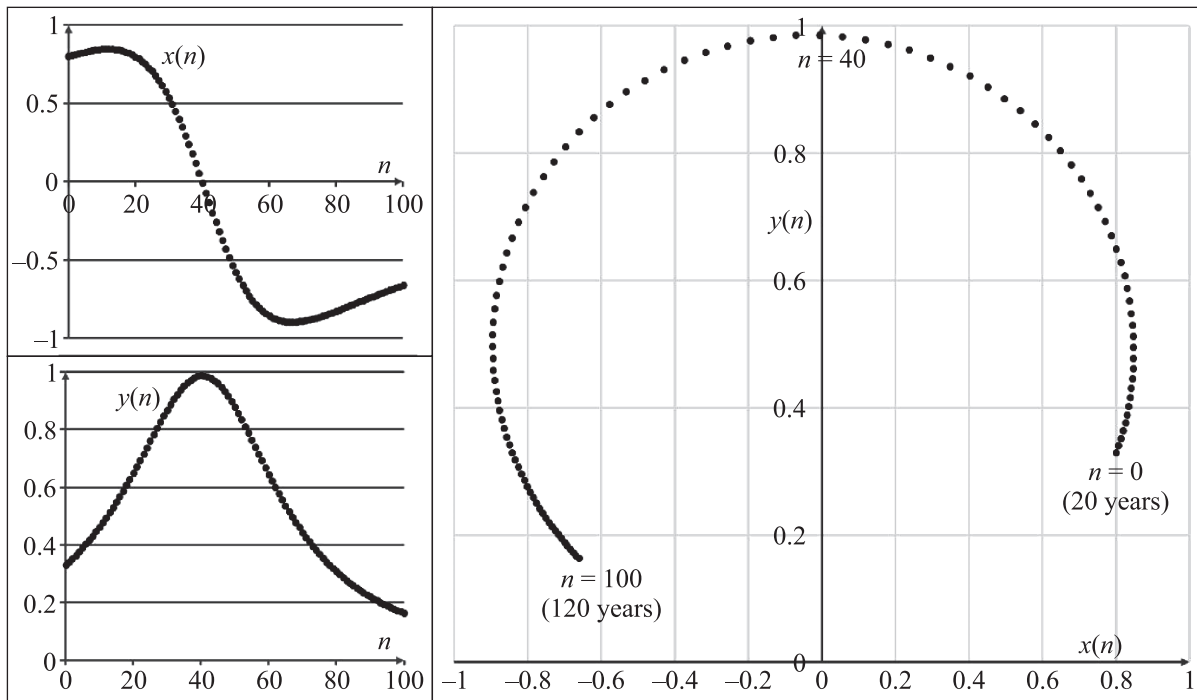


Figure 3. Cognitive abilities $x_n(x_0, y_0)$ and awareness $y_n(x_0, y_0)$ dynamics in the calibrated model for optimal trajectory A

of interest: $x_{20}, x_{40}, x_{60}, y_{20}, y_{40}, y_{60}$. They are all implicitly given in equations (1), (2) by the set of starting values x_0, y_0 . Each indicator has its own set of achievable values.

Let us visualize some of the most significant results of the trajectory analysis. On Fig. 4 on the left we present the maps of functionals x_{max} and y_{max} for all the combinations of $(x_0, y_0) \in S$. As you can see, the the maximum values of awareness and cognitive abilities are closely related: the highest level of awareness are reachable for the ones who reach the maximum cognitive abilities (point A). Maximum level of awareness occurs when the initial values of cognitive ability are high and awareness is low. This is possible in countries where the General education system is not focused on the transfer of information to students, but on the maximum development of their cognitive abilities.

As you can see, there is no contradiction between optimization criteria (y_{max} and x_{max}) in the trajectory analysis. Point A (see Fig. 4) representing the A trajectory defined above is the optimal solution for this functionals.

On Fig 4 on the right we present the set

$$\{(x_{max}, y_{max}) : x_{max} = x_{max}(x_0, y_0), y_{max} = y_{max}(x_0, y_0); (x_0, y_0) \in S\}. \tag{4}$$

Qualitative interrelations of the functionals x_{max} and y_{max} has a meaningful interpretation. The amount of information, carried by an individual, gradually increases, which requires more and more psyche’s resources to orderly preserve it. This leads to a gradual decrease in cognitive abilities and, ultimately, to the incapacity of the individual.

This non-trivial effect is achieved due to the fact that high initial awareness contributes to the growth of awareness in the future not as much (the dependence is direct linear, see Fig. 4), as it prevents the cognitive abilities development (the dependence in (1) is inverse quadratic). However, as can be seen in the same figure, the negative impact of starting awareness is insignificant for individuals with low starting cognitive abilities. This feature has important implications for the tasks of the human capital’s assessment.

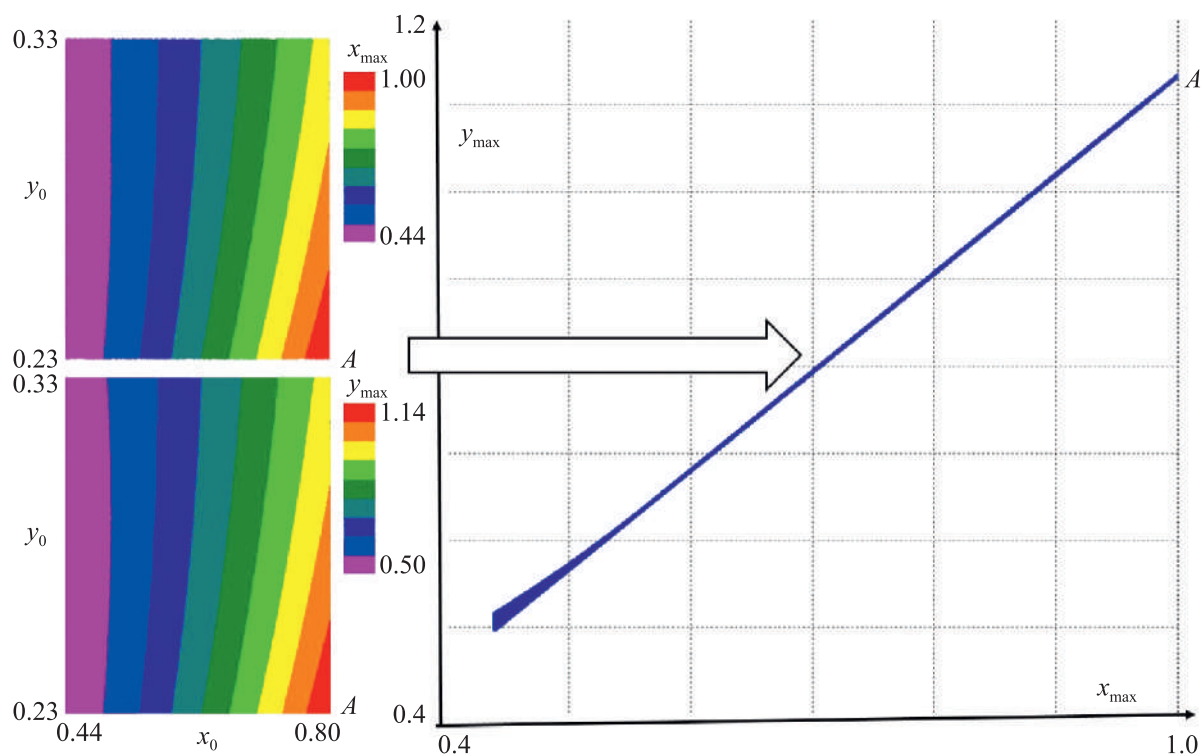


Figure 4. Sets of achievable values of cognitive abilities' maximum $x_{\max}(x_0, y_0)$ (on the left upper) and awareness $y_{\max}(x_0, y_0)$ (on the left lower) and their interrelation (on the right)

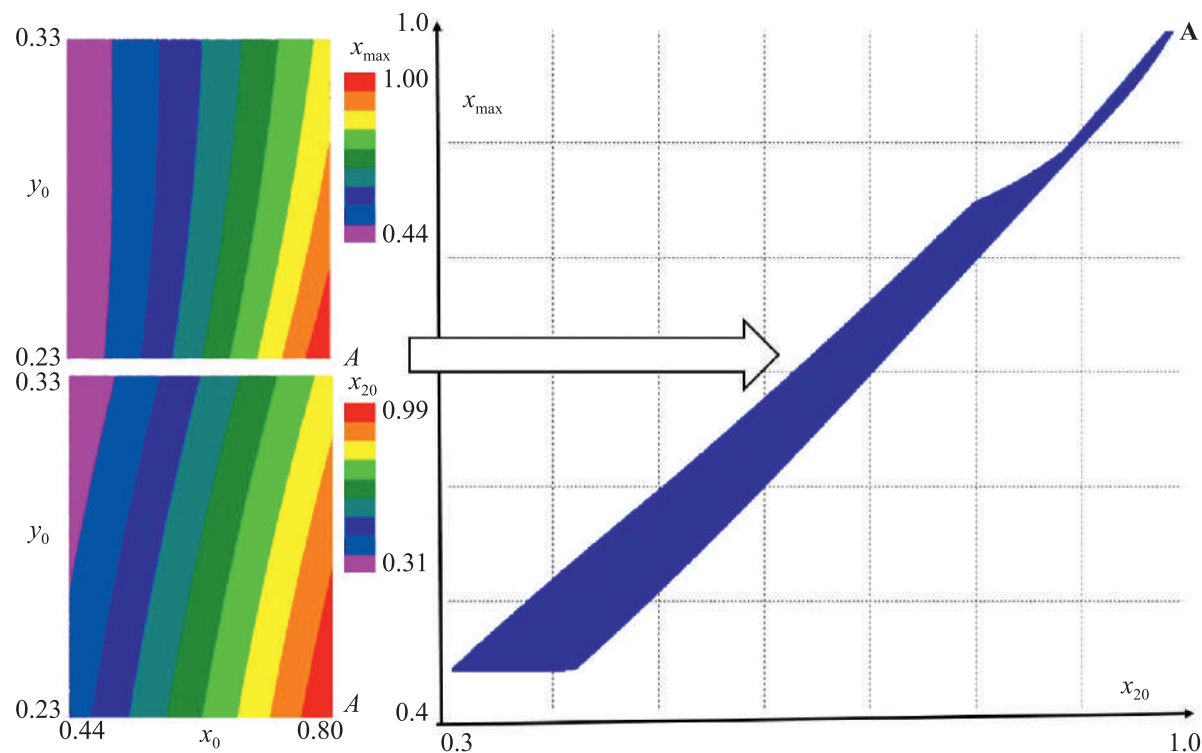


Figure 5. Sets of achievable values of cognitive abilities' maximum $x_{\max}(x_0, y_0)$ (on the left upper) and cognitive abilities by the age of $n = 20$, i.e. $x_{20}(x_0, y_0)$ (on the left lower) and their interrelation (on the right)

On Fig 5 on the right we present the set

$$\{(x_{20}, x_{\max}): x_{20} = x(x_0, y_0, 20), x_{\max} = x_{\max}(x_0, y_0); (x_0, y_0) \in S\}. \quad (5)$$

Note that highest levels of cognitive abilities are achieved in $n = 20$, but for the individuals with lower cognitive abilities x_{20} doesn't have to be the maximum (see Fig. 5). In other words, an ultra-high maximum level of cognitive ability is possible only for those individuals who have reached it already in $n = 20$ (40 years). This observation is important since it proves that despite the strict calibration conditions regarding the age of peak cognitive abilities in the population as a whole, the model allows for its large individual diversity, which is realistic.

Let's summarize the results of the primary analysis of phase trajectories. It is established that the achievement of the maximum level of awareness occurs when the initial values of cognitive abilities are high and awareness is low. Only individuals with the extremely low initial level of awareness (y_0) can reach an extremely high awareness (y_{\max}) in the model. In other words, individuals entering adulthood after primary socialization, where the main goal was the obtaining of the knowledge, will not be able to achieve the extremely high level of awareness (too large resources of their mentality are occupied with the already acquired knowledge maintaining, which prevents the further development of cognitive abilities).

Thus, the model clearly indicates the advantages of competency-based, activity-oriented, practice-oriented models of education over classical encyclopedic models in the preparation of the intellectual elite of society. This is possible in countries which education systems focus not on the transmission of information to students but on the maximum development of their cognitive abilities. At the same time, for individuals with mediocre starting cognitive abilities, a low starting level of awareness provides significantly fewer benefits.

On the other hand, it can be seen that low awareness in itself does not provide any advantage: at small y_0 , any y_{\max} values are achievable due to the influence of x_0 . The uninformed and undeveloped individual has no advantage. Reducing the initial value of awareness is useful only for individuals with high initial cognitive abilities. This effect hints to a contradiction that will be described later: how should we prioritize awareness and cognitive abilities, while that not all individuals will be able to achieve optimal starting values, and there will be a spread between individuals?

These patterns are apparent not only for the y_{\max} indicator but also for x_{\max} (this is logical given the deep interconnection of these two indicators shown on Fig. 4). Thus, starting cognitive abilities in the model determine the most achievable level of individual's development and awareness, while the starting awareness acts as a bound (non-significant for individuals with weak starting cognitive abilities since their cognitive abilities are limited anyway, and important for an individual with high starting cognitive abilities).

Multicriteria research and Pareto-frontier analysis

Let's consider the problem of individual human capital's evaluation. Classics of human capital suggest [Mincer, 1958; Becker, 1964; Schultz, 1971; Coleman, 1988] to do it basing on the lifetime return from education (school and university) acquired in childhood and youth. They use the concept of an expected return on investment, thus implementing lifelong indicators (such as expected income over a lifetime). This allowed them to formulate and solve an optimization problem: the optimal amount of investment chosen by the individual (or rather his parents) to be made in education, maximizing lifetime income.

Our model allows us to refine and detail this approach by our own method for the age dynamics of human capital' analysis. Summing up a person's cognitive abilities over the course of his life allows us to evaluate his lifetime creativity, and summing up his awareness allows us to evaluate his lifetime

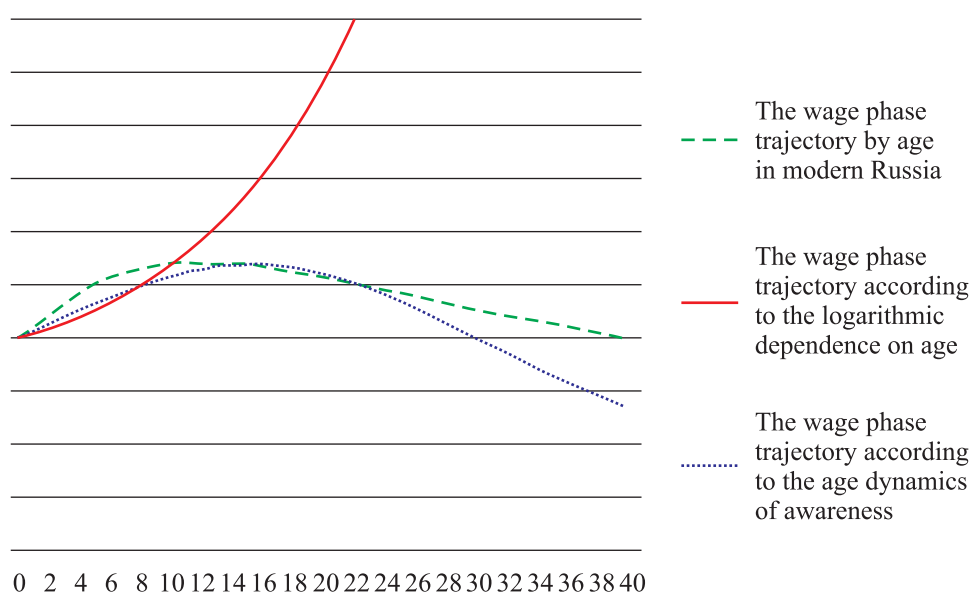


Figure 6. Typical trajectory in models of age dynamics of the Mintzer-class logarithmic model and our proposed model (years of experience in the abscissa, the relative change in the y-axis)

productivity. Please note that we do not use any discounting function. The significance of income for a person at different ages forms the independent scientific problem.

The relevance of our model can be demonstrated in a comparison of its phase trajectories with the trajectories of classical models (Mintzer's and its derivatives). The Mintzer's model was created in the 50s of the XX Century and was calibrated on data from the beginning of the XX Century. However, the expected length of employment in the United States has increased from about 27 years to 45 (with the retirement age of 65) during the first half of the 20th century. Today, when the length of working life in developed countries is limited not so much by retirement age as by healthy life expectancy (HALA), it is approaching 50. So the classical approach is poorly applicable to the modern information economy (see Fig. 6) while adequately approximating the interrelation between human capital and income under the age of 30. Additionally, the share of older workers in the employment structure continues to increase in modern information economy, and the increasing number of jobs are created in the service sector and engineering production, forming a "creative class", for which productivity (reproduction of information available to a person) is not as important as creativity (transformation of old information into new).

In contrast to the classic approach, our model introduces the missing degrees of freedom, allowing us to describe the decline in income during older ages. We did not set out model to be calibrated the by income, because it would have limited applicability this way: various factors contribute to labor income in different countries and different sectors of the economy, including cognitive abilities, awareness of the individual, his reputation, and various structural market factors. However, the very ability to describe the decline of human capital in older age is its main advantage, which opens up a great potential for its usage as a component in various models of the labor market.

However, to correctly calculate any lifetime indicators, it is necessary to consider that the probability of being alive for a person changes with his age. Generally speaking, if the peak of cognitive abilities or awareness occurs in old age, then there is a risk that the person will not have time to use them. Therefore, while summing up the considered indicators we should multiply them by this probability. So in our multi-criteria setting, any individual should maximize the amount of cognitive ability or awareness over a lifetime, weighted by the probability of survival to that age.

We will denote the probability of living to the age n as $p(n)$. Note that in our model $p(0) = 1$, because our object is adult ($n = 0$ means 20 years): meaning we assume that non-adults do not contribute to creative or reproductive production and therefore do not receive any return on their human capital. We emphasize that we are not talking about a demographic pyramid, but about the age dynamics of the population of one year of birth. Combining the model with the demographic pyramid and the birth rate forecast, it is possible to obtain an estimate of the human capital of the national economy (taking into account the change in human capital with age). However, this issue is beyond the scope of this publication, since it requires constructing of the supplementary model of the distribution of individuals by their starting values of cognitive abilities and awareness, depending on various factors, including the state’s educational policy, which differed for different generations of people who are simultaneously included in the demographic pyramid. The general approach to this supplementary problem is described in a separate publication [Kamenev, Kamenev, 2020].

The dynamics of $p(n)$ is described as

$$p_\gamma(n + 1) = (1 - \gamma k(n))p_\gamma(n), \quad \gamma k(n) \leq 1 - \delta, \tag{6}$$

$$p_\gamma(n + 1) = \delta p_\gamma(n), \quad \gamma k(n) > 1 - \delta. \tag{7}$$

Here, $k(n)$ — age mortality rates in the Russian Federation (graph 4 “mortality Tables”, see [Healthcare in Russia, 2017]), δ is a small value, equated to 0.01, and the γ is a comparative mortality rate, reflecting the demographic situation in comparison with the situation in Russia. If detailed country-specific mortality statistics are available, the country-specific mortality coefficients may be used directly instead of the value of $\gamma k(n)$.

Taking into account the probability of survival, we can construct lifelong indexes of human capital: the creativity index and the productivity index. This allows us to formulate the question in terms of optimization: what initial value of cognitive abilities and awareness maximizes the individual’s human capital? Since there are two optimal choice criteria in this problem (contradictory, as it will be shown later), we use simulation and multi-criteria optimization approaches for its study (the method of metric data analysis for model-generated data in the version of the Interactive decisions maps method [Lotov et al., 2004]. Note that the multi-criteria approach leads to not a single solution, but to the Pareto-frontier research: a set of nondominable achievable points in the criterion space (non-improvable simultaneously by all criteria) providing stakeholders with appropriate effective solutions leading to this frontier.

The general logic of our approach include approximation of the set of achievable criterion points and research of its Pareto-frontier. Approximation is performed by construction of the metric net of the criterion set. To do this, we use random samples to probe a set of combinations of initial values of the phase variables x_0 and y_0 . For each combination of initial values from the sample, we calculate the trajectory and calculate the values of the criteria functions (productivity and creativity indices). Thus, we get a two-dimensional set of achievable values of human capital indices, implicitly defined on a two-dimensional set of initial values of phase variables. This allows us to identify the Pareto-hull of the set of achievable values and research its properties.

$$x_\gamma^I(x_0, y_0) = \sum_{n=0}^{100} H(x(x_0, y_0, n))p_\gamma(n)x(x_0, y_0, n), \quad n = 0, 1, \dots, 100, \tag{8}$$

$$y_\gamma^I(x_0, y_0) = \sum_{n=0}^{100} H(y(x_0, y_0, n))p_\gamma(n)y(x_0, y_0, n), \quad n = 0, 1, \dots, 100. \tag{9}$$

Here, H is the Heaviside function, whose value is zero for negative arguments and one for positive arguments and zero, x^I — human creativity index (lifelong index, which is proportional to the total

amount of creative product that some person can create during the life, y^l – human productivity index (lifelong index, which is proportional to the total amount of (re)productive product that some person can create during the life).

Let's denote the age E , such as: $p(E)$ is as close to 0.5 as it is possible. This metric is not identical to the life expectancy, but close to it and convenient for mortality rates' calculations (we don't take into account those who doesn't survive to $n = 0$). In modern Russia $E = 54$, $\gamma = 1$, i.e. 74 years. Let's examine x^l and y^l for modern Russia. For comparison, we will consider x^l and y^l with mortality rates 100 times higher ($\gamma = 100$, $E = 5$) and 10 times less ($\gamma = 0.1$, $E = 95$). Such extreme options reflect the range of life expectancy in different countries of the modern world. This is a highly simplified model of mortality, however, it allows to draw meaningful conclusions about the impact of mortality on productivity and creativity indexes.

Let's analyze the set I_γ at different mortality rates γ (achievable bicriterial decisions) for different γ using FGM. Here,

$$I_\gamma = \{(x^l, y^l): x^l = x_\gamma^l(x_0, y_0), y^l = y_\gamma^l(x_0, y_0), (x_0, y_0) \in S\}. \tag{10}$$

First, let's plot the set I_γ for Russia (i.e. $\gamma = 1$, see Fig. 7). The point A was defined above, while the point B reflects the trajectory given by the combination of values $x_0 = x_0^{\max}$ and $y_0 = y_0^{\max}$. Here we see a weak Pareto-frontier (a set of points that cannot be improved by increasing x^l or y^l without decreasing the other one, see $[A, B]$), where the Pareto Principle comes into effect: creative capital can be significantly increased by slight sacrifice of productive capital.

Next, let's plot the set I_γ for each possible γ (see Fig. 8):

$$\{(x^l, y^l, \gamma): x^l = x_\gamma^l(x_0, y_0), y^l = y_\gamma^l(x_0, y_0), \gamma \in [0.1, 100], (x_0, y_0) \in S\}. \tag{11}$$

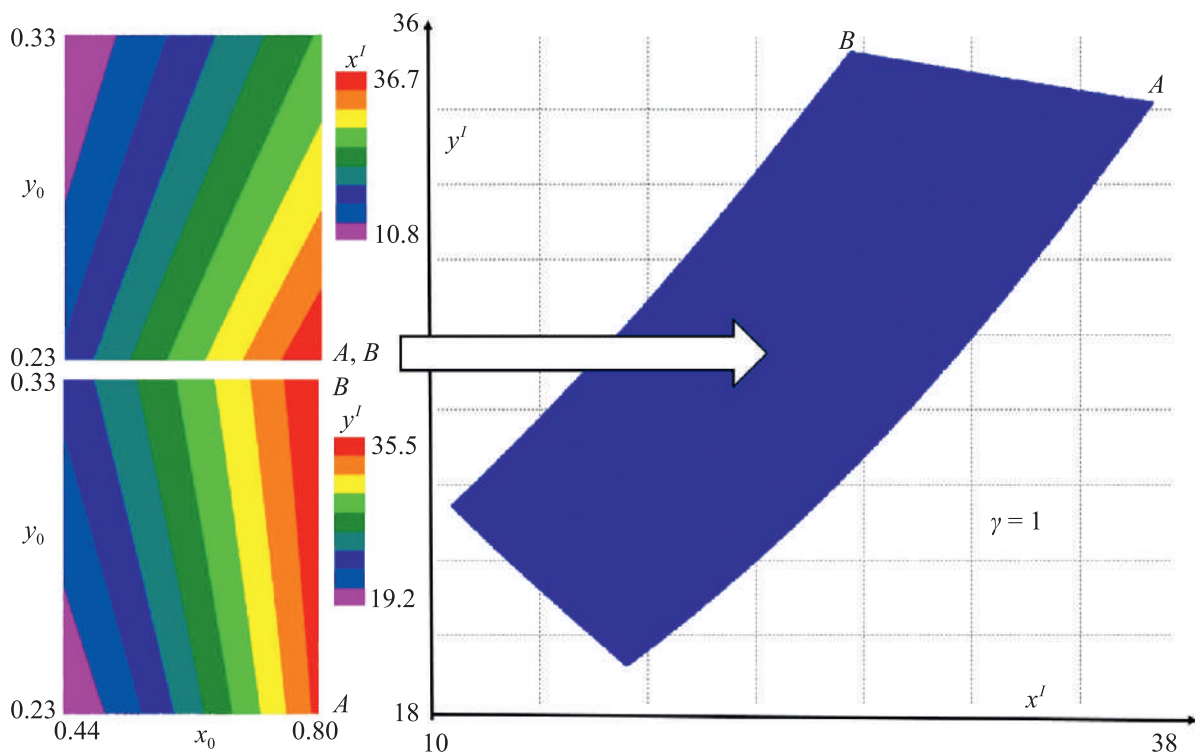


Figure 7. Sets of achievable values of creative capital's lifelong index $x^l(x_0, y_0)$ (on the left upper) and productive capital's lifelong index $y^l(x_0, y_0)$ (on the left lower) and Pareto-frontier A-B of the set I_γ of achievable values of x^l and y^l combinations for mortality rates $\gamma = 1$ typical for Russia (on the right)

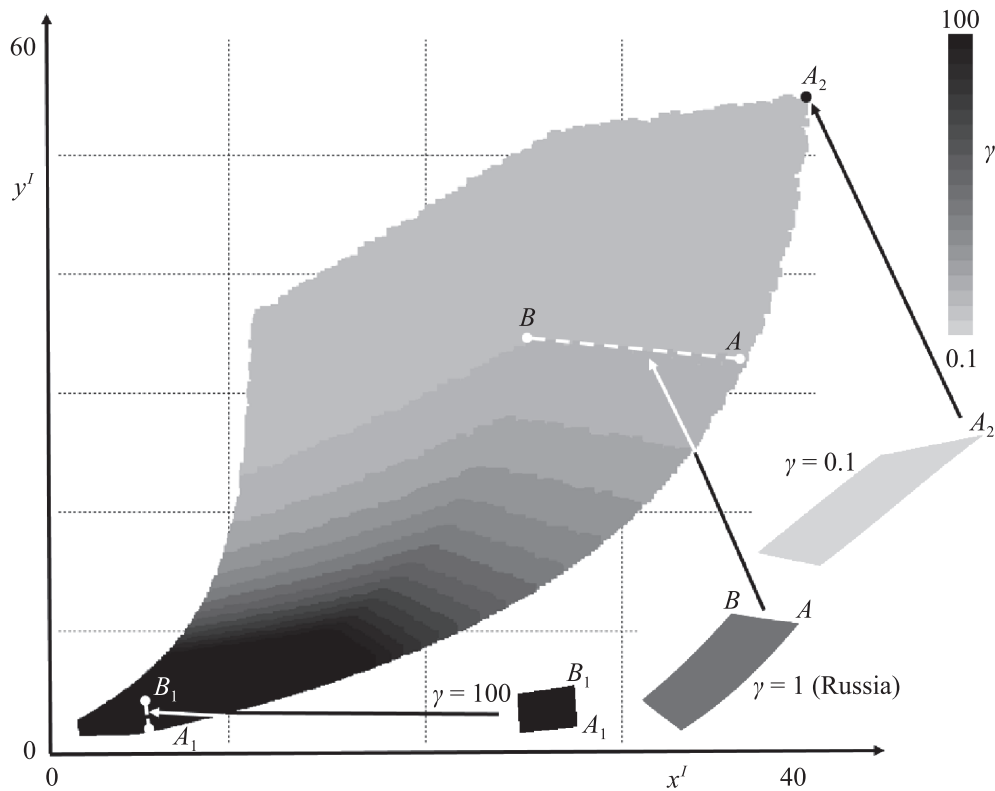


Figure 8. The evolution of Pareto-frontier $A - B$ for different mortality rates (γ)

We can see that two-factor optimization at different γ gives a different type of the Pareto-frontier. In the areas of low age-specific mortality rates we can see a weak Pareto-frontier (see $[A_1, B_1]$ for $\gamma = 100$) reversed compared to $\gamma = 1$: productive capital creative capital can be significantly increased by slight sacrifice of creative capital. With moderate mortality rates, the Pareto-frontier becomes strong, representing mutually excluded choices. This difference corresponds to the well known transition from knowledge-based education, common until the end of the twentieth Century, to a competency-based approach, appeared in the late twentieth Century and became predominant in the early twenty-first Century. For the most developed countries with high life expectancy (see A_2 for $\gamma = 0.1$), there is a dominant solution similar to the one we described earlier in the analysis of trajectories without mortality (i.e. Pareto-frontier degenerate into single point A).

Note that the $p_\gamma(n)$ function makes only a quantitative but not qualitative impact on the human creativity index x^I : even at low E , the initial awareness has a negative impact on this index, but the scale of this influence is much smaller than at extremely high E . The more life expectancy is high, the more absolute values of the index became high. The same pattern is typical for the productivity index. The influence of γ on optimization criteria is presented in more detail in Fig. 9. Here,

$$x_{\max}^I(\gamma) = \max\{x_\gamma^I(x_0, y_0) : (x_0, y_0) \in S\}, \tag{12}$$

$$y_{\max}^I(\gamma) = \max\{y_\gamma^I(x_0, y_0) : (x_0, y_0) \in S\}. \tag{13}$$

It is shown that at high γ the indexes tend to zero, and at low γ – to the theoretically possible limit, which corresponds to the optimal choice of a person who is guaranteed to live up to 120 years. Note that the influence of γ on the index of productivity is deeper, because the highest values of awareness are achieved some time after the development of maximum cognitive abilities. Therefore, the duration of the period of return on human capital is more important for the productivity index.

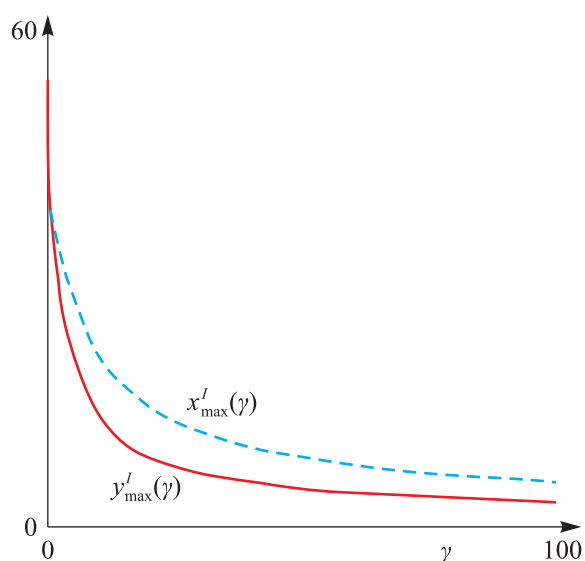


Figure 9. Maximum achievable values of the indices of creativity x_{γ}^I and productivity y_{γ}^I in population for different relative mortality rates γ

We recall that in our model it is assumed that age mortality rates vary uniformly, and childhood mortality is not taken into account. The age-specific mortality rates change can make the described dependence more complicated.

Conclusion and Significance

The study demonstrates the possibilities of metric data analysis method application for the problems of multicriteria optimization. The suggested version of the model of a person as a carrier of information takes into account two main optimization criteria: the integral characteristic of cognitive abilities and awareness of a person during the trajectory (life), weighted by the probability of survival. It is shown that these criteria can be internally contradictory (conflicting), forming in different cases a strong or weak Pareto-boundary, or a single dominant (ideal) solution. The practical significance of these boundaries is due to the fact that they describe the priorities in the individual's educational choice. It is shown that with the increase in life expectancy (decrease in mortality coefficients), the development of cognitive abilities becomes the only priority, while for societies with moderate and especially low life expectancy there is a contradictory choice between the development of abilities or absorption of specific knowledge. Russia today, as it was shown in the study, is in the conditions of the weak Pareto-boundary, when education aimed at the knowledge transfer still makes sense, but due to Pareto Principle competence-based education is preferable.

The application of the method of Metric data analysis for model identification by incomplete, fragmentary social data is also demonstrated. In this case, the known boundary conditions for the general population are considered in order to construct the calibration identification set. This approach allows to eliminate the existing gap in Social Sciences between theoretical and applied (sociometric, econometric) models, identifying phase variables including not explicitly observed.

The model can be applied mainly in refinement of the national economies human capital's estimation, as well as in the modelling of the choice of priorities of state educational policy (this problem is discussed in a separate publication, see [Kamenev, Kamenev, 2020]). For this purpose it is required to build an auxiliary model of formation of cognitive abilities and awareness in the process of education and socialization of non-adults, which is the subject of a separate article.

Список литературы (References)

- Anderson J. R., Bower G. H.* Human associative memory. — Psychology press, 2014.
- Averell L., Heathcote A.* The shape of the forgetting curve and the fate of memories // *Journal of mathematical psychology.* — 2011. — Vol. 55, No. 1. — P. 25–35.
- Barrow J. D., Tipler F. J.* The Anthropic Cosmological Principle. — Oxford Univ., 1986.
- Becker G.* Human capital: theoretical and empirical analysis, with special reference to education. — N.Y., 1964.
- Bourdieu P.* The forms of capital. — NW Biggart, 1986.
- Coleman J. S.* Social Capital in the Creation of Human Capital // *American Journal of Sociology Supplement.* — 1988. — Vol. 94. — P. 95–120.
- Crook T. R. et al.* Does human capital matter? A meta-analysis of the relationship between human capital and firm performance // *Journal of applied psychology.* — 2011. — Vol. 96, No. 3. — P. 443–456.
- Currie J., Almond D.* Human capital development before age five // *Handbook of labor economics.* — 2011. — No. 4. — P. 1315–1486.
- Davies P. et al.* Paying for education: debating the price of progress. — Routledge, 2018.
- Fleisher B., Li H., Zhao M. Q.* Human capital, economic growth, and regional inequality in China // *Journal of development economics.* — 2010. — Vol. 92, No. 2. — P. 215–231.
- Frederick S.* Cognitive reflection and decision making // *Journal of Economic perspectives.* — 2005. — Vol. 19, No. 4. — P. 25–42.
- Galor O., Weil D. N.* Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond // *American economic review.* — 2000. — Vol. 90, No. 4. — P. 806–828.
- Galor O., Weil D. N., de la Croix D., Fiaschi D., Fioroni T., Dalgaard C. J., Strulik H., Doepke M., Kindermann F., Cinnirella F., Hornung E.* Demographic Change and Long-Run Development. — MIT Press, 2017.
- Gimpelson V., Kapeliushnikov R.* Between light and shadow: Informality in the Russian labour market // *The Challenges for Russia's Politicized Economic System.* Routledge. — 2015. — P. 57–82.
- Gimpelson V. E.* Does the Russian economy need human capital? Ten of the doubt [Nuzhen li rossijskoj ekonomike chelovecheskij kapital? Desyat somnenij] // *Voprosy ekonomiki.* — 2016. — Vol. 10. — P. 129–143.
- Gimpelson V. E., Kapelyushnikov R. I., Lukyanova A. L.* Demand for labor and qualification in industry: between deficit and overstocking [Spros na trud i kvalifikaciyu v promyshlennosti: mezhdru deficitom i izbytkom] // *Ekonomicheskij zhurnal Vyshej shkoly ekonomiki.* — 2007. — Vol. 2, No. 11. — P. 163–199.
- Gimpelson V. E., Kapelyushnikov R. I., Sharunina A. V.* The Roads of our choice: changes in the external and internal labor markets [Dorogi, kotorye my vybiraem: peremeshcheniya na vneshnem i vnutrennem rynkah truda] // *Ekonomicheskij zhurnal Vyshej shkoly ekonomiki.* — 2016. — Vol. 2, No. 20. — P. 201–242.
- Healthcare in Russia [Zdravoohranenie-v-Rossii-2017], Statistical review. — Rosstat, 2017.
- Henderson J.* Memory and oblivion. — Psychology Press, 1999.

- Kamenev G. K., Lyulyakin O. P., Sarancha D. A., Lysenko N. A., Polyansky V. O.* From chaos to order. Difference equations in one ecological problem // Russian Journal of Numerical Analysis and Mathematical Modelling. — 2016. — Vol. 31, No. 5. — P. 253–265.
- Kamenev G. K., Olenov N. N.* Study of the Russian economy's identification and forecast stability using a Ramsey type model // Mathematical Models and Computer Simulations. — 2015. — Vol. 7. — P. 179–189.
- Kamenev I. G.* Modeling of a person as a carrier of information using the method of Metric Data Analysis [Modelirovanie cheloveka kak nositelya informacii s primeneniem metoda metriceskogo analiza dannyh] // Proceedings of the IX Moscow conference of operations research ORM2018 [Trudy IX Moskovskoj mezhd. konf. po issledovaniyu operacij (ORM2018)]. Part 2. — 2018. — P. 179–184.
- Kamenev G. K., Kamenev I. G.* Metric analysis of multidimensional sociological samples [Metriceskij analiz mnogomernyh sociologicheskikh vyborok] // Pospelov I. G. et al. (eds.) Proceedings of the conference “Modeling of coevolution of nature and society: problems and experience. To the 100-th anniversary from the birthday of academician N. N. Moiseev” [Modelirovanie koevolucii prirody i obshchestva: problemy i opyt. Trudy Vserossijskoj nauchnoj konferencii]. — Moscow: FITZ of RAS, 2017. — P. 198–209.
- Kamenev G. K.* A Visual Method for Parameter Estimation // Doklady Mathematics. — 1998. — Vol. 57, No. N2. — P. 305–307.
- Kamenev G. K.* Approximation of Completely Bounded Sets by the Deep Holes Method // Computational Mathematics and Mathematical Physics. — 2001. — Vol. 41, No. 11. — P. 1667–1675.
- Kamenev G. K.* Dynamic model calibration with conditions on the tube of trajectories [Kalibrovka dinamiceskikh modelej s usloviyami na trubku traektorij] // *Ereshko F.* (editor-in-chief) Proceedings of IX Moscow international Conference on Operations Research, 2018 [Trudy IX Moskovskoj mezhd. konf. po issledovaniyu operacij (ORM2018)] — Germeyer100 (ORM2018, Moscow, 22–27 oct. 2018). — Moscow: MAXS Press, 2018. — Vol. 2. — P. 360–364.
- Kamenev G. K.* Multicriteria Identification Sets Method // Computational Mathematics and Mathematical Physics. — 2016. — Vol. 56, No. 11. — P. 1843–1858.
- Kamenev G. K., Kamenev I. G.* Discrete-dynamic modeling of governance for human capital [Diskretno-dinamicheskoe modelirovanie gosudarstvennogo regulirovaniya chelovecheskogo kapitala] // Mathematical Modelling [Matematicheskoe modelirovanie]. — 2020. — Vol. 32, No. 6. — P. 81–96.
- Kamenev G. K., Kamenev I. G.* Multidimensional statistical sets and their metric analysis // Proceedings of the Institute for System Analysis of the Russian Academy of Sciences CC FRC of CSC of the RAS. — 2018. — Vol. 68, No. 2. — P. 30–33.
- Kamenev G. K., Kamenev I. G.* The model of human as information carrier and its metric data analysis. In: *Matiushka V. M., Lasanuck I. V.* (editors) Conference proceedings of VII International Conference “New trends, strategies and structural changes in emerging markets” (Moscow, May 29–31, 2018). — Moscow: RUDN, 2018. — P. 67–69.
- Kapelyushnikov R. I.* Demand and supply of highly skilled labor in Russia: who ran faster [Zapiska ob otechestvennom chelovecheskom kapitale]. — Moscow: GU VSHE, 2011.
- Kapelyushnikov R. I.* How much is the human capital of Russia? [Skolko stoit chelovecheskij kapital Rossii?]. — Moscow: GU VSHE, 2012.
- Kapelyushnikov R. I.* Note on domestic human capital [Zapiska ob otechestvennom chelovecheskom kapitale]. — Moscow: GU VSHE, 2008.

- Kapelyushnikov R. I., Lukyanova A. L.* Transformation of human capital in Russian society (basing of “Russian monitoring of economic situation and population health”) [Transformaciya chelovecheskogo kapitala v rossijskom obshchestve (na baze “Rossijskogo monitoringa ekonomicheskogo polozheniya i zdorovya naseleniya”)]. — Moscow: Liberalnaya missiya, 2010.
- Kolmogorov A. N., Tikhomirov V. M.* Epsilon-entropy and Epsilon-capacity of sets in functional spaces // *Advances in mathematical Sciences*. — 1959. — Vol. 25, No. 2. — P. 3–86.
- Korickij A. V., Grigorova T. V.* Human capital as a factor of economic growth of Russian regions [Chelovecheskij kapital kak faktor ekonomicheskogo rosta regionov Rossii]. — Novosibirsk: Sibirskij universitet potrebitelskoj kooperacii, 2010.
- Lotov A. V., Bushenkov V. A., Kamenev G. K.* Interactive Decision Maps. Approximation and Visualization of Pareto Frontier, Applied Optimization **89** Kluwer Academic Publishers. Boston / Dordrecht / New York / London, 2004.
- Mincer J.* Human Capital Responses to Technological Change in the Labor Market. — Cambridge, 1989.
- Mincer J.* Investment in Human Capital and Personal Income Distribution // *Journal of Political Economy*. — 1958. — Vol. 66, No. 4. — P. 281–302.
- Ployhart R. E., Moliterno T. P.* Emergence of the human capital resource: A multilevel model // *Academy of management review*. — 2011. — Vol. 36, No. 1. — P. 127–150.
- Radaev V. V.* The concept of capital, forms of capital and their conversion [Ponyatie kapitala, formy kapitalov i ih konvertaciya] // *Ekonomicheskaya sociologiya*. — 2002. — Vol. 4, No. 3. — P. 20–32.
- Salthouse T. A.* Within-Cohort Age-Related Differences in Cognitive Functioning // *Psychological Science*. — 2013. — Vol. 24, No. 2. — P. 123–130.
- Schultz T.* Investment in human capital: the role of education and of research. — N.Y., 1971.
- Shabunova A. A., Leonidova G. V.* Human capital: indicator of sustainable development of the territory [Chelovecheskij kapital: indikator ustojchivogo razvitiya territorii] // *Ekonomicheskie i socialnye peremeny: fakty, tendencii, prognoz*. — 2011. — Vol. 5, No. 17. — P. 101–115.
- Shannon C.* The Mathematical Theory of Communication // *The Bell System Technical Journal*. — 1948. — Vol. 27, No. 3, 4. — P. 379–423, 623–656.
- Smirnov V. T., Soshnikov I. V., Romanchin V. I., Skoblyakova I. V.* Human capital: the content and types, assessment and stimulation [Chelovecheskij kapital: sodержanie i vidy, ocenka i stimulirovanie] / edit. *Smirnov V. T.* — Moscow: Mashinostroenie-1, Orel: OrelGTU, 2005.
- Upravlenie personalom i chelovecheskij kapital sovremennoj Rossii: kollektivnaya monografiya [Personnel management and human capital in modern Russia: the collective monograph] / eds. *O. I. Geliga, V. P. Solomin, G. L. Tulchinsky*. — St. Petersburg: Knizhnyj Dom, 2011.
- Zabelina O. V., Kozlova T. M., Romanyuk A. V.* Comparative assessment of human capital of Federal districts of the Russian Federation [Srvnitelnaya ocenka chelovecheskogo kapitala federalnyh okrugov Rossijskoj Federacii] // *Regionalnaya ekonomika: teoriya i praktika*. — 2014. — Vol. 37. — P. 2–9.

