National Research University
Higher School of Economics

As a manuscript

Gorbunova Elena Vasilyevna

Development of a methodological approach to comparing factors of students attrition from universities in Russia and the USA

Thesis summary for the purpose of obtaining PhD in Sociology

Academic Supervisor:

Doctor of Sciences,

Professor Ulyanov Vladimir V.

Moscow, 2024

This manuscript was prepared at Faculty of Social Science, School of Sociology.

Main results of study are presented in publications:

1. Kondratjeva O., Gorbunova E. and Hawley J. D. Academic Momentum and Undergraduate Student Attrition: Comparative Analysis in US and Russian Universities // Comparative Education Review. - 2017. - №3. - P. 607-633.

2. Gorbunova E.V. and Ulyanov V.V. Diskretnyye modeli analiza nastupleniya sobytiy: razrabotka podkhodov k sovmeshcheniyu dannykh, imeyushchikh raznuyu periodichnost' [Discrete models of Event History Analysis: elaboration of approaches to combining data with different periodicities] // Sotsiologiya: metodologiya, metody, matematicheskoye modelirovaniye [Sociology: Methodology, Methods, Mathematic Modelling]. 2016. - № 43. - P. 128-153.

3. Gorbunova E.V., Ulyanov V.V. and Furmanov K. K. Using data from universities with different structure of academic year to model student attrition // Applied Econometrics. 2017. - № 45. - P. 116-135.

4. Gorbunova E.V. Student attrition from universities: research in Russia and the United States of America // Educational Studies. 2018. - №1. - P.110-131.

Selected conferences:

- 31th CHER conference (Consortium of Higher Education Researchers) «Differentiation and Integration in Higher Education: Patterns and Dynamics» (Moscow, 2018). Presentation: «Exploring factors of undergraduate student departure in Russian and the U.S. universities».
- 8th International Academic Conference for Students and Graduate Students ‘Statistical Methods for Analysis of the Economy and Society’ (Moscow, 2017). Presentation: «Building the Model of the Students’ Withdrawal by the University Data with the Different Midterm Examination Periodicity».
- VIII Moscow International Conference of Higher Education Researchers "Dynamic Universities for Dynamic World" (Moscow, 2017). Presentation: «Building the Model of the Students’ Withdrawal by the University Data with the Different Midterm Examination Periodicity»
- 7th International Academic Conference for Students and Graduate Students ‘Statistical Methods for Analysis of the Economy and Society’ (Moscow, 2016). Presentation: “Comparison of approaches for combining data with Different Midterm Examination Periodicity in Event History Analysis”.
- VII Moscow International Conference of Higher Education Researchers "University between Global Challenges and Local Commitments" (Moscow, 2016). Presentation: «Exploring Factors of Undergraduate Student Departure in Russian and US Universities».
- Comparative and International Education Society (CIES) conference “Ubuntu! Imagining a humanist education globally” (Washington, DC, 2015). Presentation:

«Exploring factors of undergraduate student departure in Economics and Engineering discipline in U.S. and Russian universities».

- IV International Conference of the Russian Association of Higher Education Researchers (Moscow, 2013). Presentation: “Comparative Analysis of Undergraduate Student Attrition in US and Russian Universities”.

Introduction (problem statement, degree of development of the problem)

When conducting comparative social research, the researcher faces a number of issues, such as the selection of objects for analysis, theoretical framework, operationalization of research concepts, selection of relevant data, etc. One of the important methodological tasks is to ensure data coherence and the possibility of joint analysis. Often there is a situation when data on compared objects, be it individuals, institutions, countries, are carried in different formats, which creates difficulties in the analysis.

Likewise, in this paper, devoted to the analysis of factors of students leaving universities in Russia and the USA, the problem of sharing data entered into the system in different formats has been solving. The subject of the empirical analysis was the factors of student attrition, and the dependent variable was both the fact and the time of attrition. In the US universities, periods in which student expulsion was recorded differed - in some universities they were semesters (the duration of one semester is 4.5 months), and in others - trimesters (the duration of one trimester is approximately three months). Since a prerequisite for the use of event history analysis models (hereinafter referred to as EHA) in comparative analysis is a single dimension of the time measure [Allison, 1984], it became necessary to develop a way that would combine different periodicities in one model. An additional parameter of this task was the inclusion of time-varying covariates in the model. As far as the authors know, the problem of combining different periodicities has not previously been solved within the framework of the EHA.

The development of a methodology for combining data with different periodicities in the EHA, using the example of a comparative analysis of data on the departure of university students in Russia and the USA, became the central element of this work. Important results of the work were also the testing of the theoretical concept of “academic momentum” in a comparative study of student attrition and drawing conclusions about the factors influencing their expulsion. This theory was first applied to statistical data, as well as in a comparative study of student dropout from universities.

The problem of student expulsion from university has been an important aspect of study in the field of students research as well as university administration for a long time [Braxton et al., 2013; Manyanga, Sithole and Hanson, 2017]. On the one hand, it has a number of economic and social negative consequences for students, the state, and households [Tinto, 1987; Pascarella and Terenzini 1991; DesJardins, Ahlburg and McCall, 1999; Kehm, Larsen and Sommersel, 2019]. On the other hand, the process of student expulsion is an important mechanism for controlling the quality of student learning, without which education will not fulfill the function of a social elevator [Gruzdev, Gorbunova and Frumin, 2013]. In some cases, expulsion does not mean a loss for the student or the education system, but it is only a transfer to another university, to another specialty. Often after a break in education (which is not considered as a loss, but as an enrichment of his/her experience on the contrary), a student returns to the educational system and successfully completes his/her studies. However, often university loses students whom it should not lose - the reason is not because of their low level of training or conscious refusal to receive a particular education, but because of the complexity of their social adaptation, difficulties of interaction with a professor,

lack of services that can help the student to cope with emerging difficulties in the educational process.

Today it is becoming relevant for Russian universities to explore how to keep the contingent of students without reducing the quality of the educational process. At the same time, the few works in this research field do not provide a full-fledged picture of the scale of departure and factors affecting student dropout have a weak theoretical and methodological basis [Gorbunova, 2018]. It becomes important to get an answer to the questions concerning universal factors that influence students' withdrawal, as well as specific factors with regard to particular groups of students, types of higher education institutions. It is known that in comparison with other countries the dropout rate in Russian universities is not the highest [OECD, 2010], however, there are no researches examining the differences between the process and factors of student dropout in Russian universities and dropout in other countries. It is possible to get answers to these questions with the help of international and inter-institutional comparative studies.

While conducting comparative researches, an important aspect becomes the development of methodology, including both the choice of theoretical framework, object of research, data sources, and methodology of performing comparative analysis. Our study is focused to these aspects exactly – the selection and validation of the theoretical framework, the construction of the research design itself, as well as the methods of comparative analysis under conditions of different data formats. This task will be performed on the data on student expulsion from Russian and American universities.

Comparative studies provide an opportunity to study how the phenomenon of interest to us is explained by the influence of global and national trends, local factors associated with the cultural and historical development of the country [Marginson and Rhoades, 2002]. Anglo-Saxon countries such as the USA and England are often used as an object of comparison in higher education, which are considered as the “gold standard” of the educational system [Kosmützky and Nokkala, 2014]. And in this study, for comparison with Russia, we chose a country in which the study of student attrition from universities has received the greatest development (research on attrition from American universities has been conducted since the 1930s [Gorbunova, 2018]), and the development of measures to reduce attrition is important task of higher education institutions policy. In addition, despite existing cultural differences and differences in the degree of freedom provided to students and requirements for academic performance [Kuzminov and Yudkevich, 2007], the educational systems of Russia and the United States are massive and have much in common in historical and institutional development. Moreover, with the globalization of higher education, the educational process at universities in both countries are becoming similar (introduction of adaptation courses in several Russian universities, transition to a bachelor's-master's degree system, individualization of the curriculum).

Purpose and tasks of study

The purpose of this study is to develop a methodology of combining data of students' expulsion with different periodicity within the framework of discrete EHA models for analyzing the occurrence of events with dynamic covariates and testing them on data on student departures from Russian and American universities.

Object of study (methodological) – techniques for combining data with different periodicities within the framework of discrete EHA models for analyzing the occurrence of events with dynamic covariates.

The object of the study (empirical) is university students in Russia and the USA.

The subject of the study is the possibilities and limitations of methods for combining data with different periodicities within EHA models for analyzing the occurrence of events using the example of studying the factors of students leaving universities in Russia and the USA.

The tasks of our research were as follows:

1) to propose a theoretical model for analyzing the factors of student attrition in American and Russian universities, taking into account the characteristics of national educational systems;

2) to develop methods for combining data with different periodicities within the framework of discrete EHA models for analyzing the occurrence of events with dynamic covariates;

3) to explore the possibilities and limitations of the proposed methods for combining data with different periodicities within the framework of discrete EHA models for analyzing the occurrence of events with dynamic covariates using the example of analyzing student attrition factors in Russian and American universities and create recommendations for their use;

4) to test the theoretical model of "academic momentum" as a part of a comparative analysis of the factors of student attrition in American and Russian universities, describe the factors of student attrition in Russian and American universities as a result of testing the proposed theoretical model.

The development of such techniques is relevant, given the increasing use of administrative data in the study of students' experience at university, as well as the increasing use of non-reactive data in sociological research [Saponova and Kulikov, 2021]. EHA methods are also often applied in demographic studies [Denisova, 2010; Tyndik and Biryukova 2015; Artamonova and Mitrofanova, 2016; Mitrofanova 2020], labor market studies [Denisova, 2017; Sinyavskaya, Chervyakova and Horvat, 2022], where time variable may not be accurately recorded in the data. Individuals may not remember the specific date of an event (for example, returning to work or becoming pregnant) and indicate it in lower frequency - month or time of year. In this case, the researcher is faced with the fact that data for the same variable are entered on different scales. And he is faced with the task of taking this difference into account when conducting the analysis. He can either reduce the data to a single periodicity (for example, interpolate to dates and treat the process as continuous, or aggregate to a lower frequency and treat the process as discrete), or use methods that do not require reduction to a single periodicity and successfully cope with data analysis at different intervals. Here are some examples of such studies. One of them is the Russian Longitudinal Monitoring Survey[1], in which study participants indicate their date of birth in different ways - as an exact date, or a month and year, or only a year. Others are sociological studies "Person, family, society"[2] (for example, the variable of the date of birth of children is entered in different formats), "Parents and children, men and women in the family and society"[3]

[1] https://www.hse.ru/rlms/

[2] https://social.ranepa.ru/tsentry-i-instituty/institut-sotsialnogo-analiza-i-prognozirovaniya/issledovaniya/86-chelovek-semya-obshchestvo-2017

[3] https://www.hse.ru/org/hse/4432173/mathbase/databases/db_11

(for example, the date of start of residence in a locality is indicated either in the periodicity of time year, or a specific month). In these and other sociological studies and data, the need arises for the most effective strategy to combine different periodicities in the EHA, and the development of an appropriate methodology is relevant.

The main task of the study - the development of methods for combining data with different periodicities in EHA - will be accompanied by the task of testing these techniques on data from a comparative study of student departures from Russian and American universities. In the course of solving this task, it is planned to obtain conclusions about the suitability of each methodology for solving the problem of combining different periodicities of data in EHA, as well as to describe the criteria for selecting a particular methodology depending on the research objectives and data.

The development of methodologies for combining data with different periodicities in EHA is the centerpiece of this study; however, the study covers a broader aspect of developing a methodology for conducting benchmarking, which also includes the selection of the object of study, data sources, methods for conducting the study, and approbation of the theoretical framework. The study will test the theory of "academic momentum", focusing on the characteristics of the educational process. This theory will be piloted for the first time with Russian data, as well as within the framework of a comparative research of students' dropout from universities.

Thus, the present study contributes both to the development of techniques for combining data with different periodicities within the framework of analyzing discrete processes of event occurrence, and to the development of methodology for conducting comparative studies, testing the theoretical framework of "academic momentum" and studying the factors that influence student expulsion from universities in Russia and the United States.

Theoretical basis of research

In this section, we describe both the history of the development of the methodology of combining particulars with different periodicities and the theories of departures with special focus on "academic momentum" theory that was used as a theoretical framework for a comparative study of student attrition between Russian and American universities.

Evolution of methods for combining data with different periodicities

The focus of our study is a variable that fixes the time parameter of some phenomenon. Time can be fixed as a continuous variable (with the accuracy of a day, a minute, a second) or as a discrete variable (with the accuracy of a certain observation interval). The situation often occurs where temporal data are fixing in discrete variables with different scale of measurement. In one case, the measurement scale for the same variable is the same, but the measurement scales are different for the variables included in the model. This situation is often encountered in GDP (gross domestic product) surveys when some variables are registered once a quarter, while others are fixed on a monthly or daily basis. In another case, there are situations when the same variable is recorded with different periodicity. One such example is the combination of data on the variable that captures the fact of student dropout in US universities, where in some universities this event was recorded in semesters, in others - in trimesters [Chiang, 2012].

In econometric studies, when building models by means of time series analysis based on data with different periodicity, researchers most often apply the procedure of data aggregation with reduction to the mean [Foroni and Marcellino, 2013]. There are also more complicated methods, such as state-space models [Evans, 2005], Mixed data sampling models [Ghysels, 2006], Mixed frequency vector autoregressive models [Ghysels, 2016] and others.

In analysis of the occurrence of events, which is considered in this dissertation study, the phenomenon when time is not recorded continuously but in certain intervals (month, semester, year, etc.) is denoted as interval censoring [Allison, 1984]. One of the main conditions for the application of EHA models is the uniform dimensionality of the time parameters included in the analysis. In the situation of working with discrete data containing information about the time of occurrence of an event, when there is not just interval censoring, but the length of intervals for different observations is unequal, or the intervals overlap, the problem of data matching arises. This problem can be solved by missing data imputation, when a certain value is assigned to the missing parameter. The most commonly used procedure is data reduction to the mean, but this procedure is also not without drawbacks, e.g., shift of parameter estimates at large intervals, underestimation of standard errors [Kim, 2003; Law and Brookmeyer, 1992]. On the other hand, the procedure of data aggregation to lower frequencies is used. So, in research [Chiang, 2012] data on the dependent variable - time of student dropout - for universities with trimester and semester systems of education were combined, and a method of aggregating data to lower frequencies, in particular, to annual intervals was used.

According to the results of our analysis, there is no distinct line of research in the literature focusing on combining data with different periodicities in EHA. There may be several reasons for this. On the one hand, EHA models are more often formulated in continuous time. On the other hand, it is a rare case when the process is discrete in nature (as in the case of student expulsion - which is usually linked to the period of a session or the beginning of a new period of study). Finally, the use of EHA models has so far been predominantly based on data where the problem of different periodicity does not arise (using survey data, or non-response data collected according to a single design).

At the same time, today there is an increase in the use of non-reactive sources, "digital footprints", big data, in particular, in the field of higher education research, when data are collected for several universities with different periodicity of the academic year, so the development of this methodological direction is seen as perspective.

The theory of "academic momentum". The state of research on student attrition from Russian universities

The study of methods for combining data with different periodicities was carried out using the example of a comparative study of the factors of student departure from universities in Russia and the USA. To carry out the empirical analysis, the theory of “academic impulse” was chosen. This section will provide a brief excursion into this and other theories of student attrition, as well as the rationale for choosing this particular theoretical framework.

It is worth noting that the history of studying student attrition goes back more than 80 years. To date, the main theories of attrition have been developed in the fields of economics, sociology, psychology, and organizational science [Habley, Bloom and Robbins, 2012].

Sociological theories view student attrition in terms of social structures and forces, taking into account the student's environment at the university, family status, early socialization mechanisms, and support from significant others. Among the main sociological concepts, it is worth mentioning the theory of J. Berger, who analyzes the process of retirement based on the concept of cultural capital [Berger, 2000], the cultural theory of G. Kuh and P. Love [Kuh and Love, 2000], and the theory of integration by V. Tinto [Tinto, 1975; 1993].

The concept of integration [Tinto, 1975] occupies an almost paradigmatic status in the field of research on student dropout from universities. This theoretical framework was proposed by V. Tinto in 1975, and was later refined by the author himself and other researchers. Considering the university as a social system with a value and social structure, Tinto noted that the main factors for student departure are insufficient integration into the social and academic systems. The central element of integration theory is the process of interaction of the individual with the academic and social systems of the university, during which the goals and institutional obligations of the individual change, which ultimately leads to various forms of attrition.

It is worth noting that this theory focuses on the attrition of students from the higher education system (not the university), and the decision to attrition is considered as the dependent variable. Considering these factors, the theory of integration is not suitable for the purposes of the comparative analysis of this work, since departure from Russian universities in most cases is forced rather than voluntary (most often the prevailing reason is "academic failure"). In addition, the data used for analysis does not allow us to track whether the student remained in the higher education system after leaving (moved to another university) or not.

Psychological concepts of attrition take into account characteristics and processes at both the individual and environmental levels, such as self-confidence, resistance to stress, high efficiency, internal locus of control [Bean and Eaton, 2001], and student motivation [Deci and Ryan, 1991], etc. [Demetriou and Schmitz-Sciborski, 2011].

Economics considers retirement primarily within the framework of human capital theory [Becker, 1964]. Attrition occurs when a student believes that the costs of attending a particular institution begin to outweigh the benefits. Costs include tuition fees, the potential income that the student would receive if he worked, benefits include future wages, gaining additional skills, knowledge, and life satisfaction.

It is worth noting that an important factor in the choice of theoretical framework was the characteristics of the data used. In particular, the analysis was based on administrative data from universities, which did not contain information about the features of student integration, psychological characteristics, life satisfaction, or assessment of acquired skills. The study of methods for combining data with different periodicity was carried out on the example of a comparative study of the factors of student attrition from universities in Russia and the United States. The theory of "academic momentum" was chosen for the analysis.

As part of a comparative study of attrition factors in Russian and American universities, the theory of "academic momentum", developed by K. Adelman [Adelman, 1999, 2006] and P. Attewell [Attewell, Heil and Reisel, 2012], was used. This conceptual framework focuses on the characteristics of the educational process and is well suited for the analysis of administrative data.

The concept of "academic momentum" was first proposed by K. Adelman. The concept of "academic momentum" was first proposed by K. Adelman in 1999 and 2006. [Adelman, 1999; 2006]. Based on results of the National Education Longitudinal Study, K. Adelman found that students who move at a higher "speed" are more likely to successfully complete their studies, compared to students who move slowly or interrupt their studies. K. Adelman introduced the concept of "academic momentum" to denote this velocity of learning, and in his study he considered such indicators of this concept as intensity of study in high school, break between graduation from high school and enrollment in higher education, academic performance in the first year of higher education, changes in academic performance at university, inclusion of adaptation courses in the student's curriculum, employment during higher education, absence of breaks in higher education, studying courses in the summer period [Adelman, 1999; 2006]. Adelman's researches are inductive in nature and rather consider correlative connections, introduce the concept of academic momentum, but do not offer a theoretical basis for considering this phenomenon.

At the Attewell's and colleagues research the work on this concept has been continued, a reflection on the use of indicators of "academic momentum" already proposed earlier by Adelman was made and some theoretical foundations were suggested [Attewell, Heil and Reisel, 2012]. Attewell and colleagues proposed to consider only a part of the indicators characterizing the student's efforts at university, namely, the break between graduation from high school and studying at university, the intensity of the academic load in the first semester of study at university, as well as the presence of courses in the summer period after the first year of study. In terms of the development of the theory of "academic momentum", this paper has described several mechanisms how this concept can be related to student attrition. First, a higher academic workload in the first period of higher education provides a higher degree of integration of the student into the life of the university, which, according to the dominative in this research field integration theory of V. Tinto, is related to the student's sharing of the values of the university and its decision. Tinto, is related to the student's sharing of the values of the university and his/her decision to drop out [Tinto, 1975]. Another mechanism is an explanation of how a higher learning load increases a student's self-efficacy skill and self-efficacy and thus is associated with successful graduation. This mechanism has already been described in studies [Bandura, 1997; 2001; Zimmerman, 2000]. And the third set of mechanisms is the life circumstances that prevent the student from a large academic load, such as insufficient financial aid, family responsibilities, and the availability of paid work. According to previous research, these factors have a significant relationship with student attrition [Braxton, 2000; Kuh, Kinzie, Schuh, 2010; Seidman, 2005; Tinto, 1993].

To date, the "academic momentum" theory has been tested on American data [Adelman, 1999; 2006; Attewell, Heil and Reisel, 2012; Attewell and Jang, 2013; Douglas and Attewell, 2014; Wang et al., 2015; Wang, 2015; Davidson and Blankenship, 2017; Zhang, 2019] as well as data on Australian universities [Martin et al., 2013]. These studies confirmed the key points of this theory, although the results on the influence of individual indicators differ.

The Russian field of research on student dropout is at the stage of formation, at the same time, the field of being studied topics is diverse. It analyzes the scale of attrition in different universities and in different areas of training [Rutkevich, 2002; Kolotova, 2011; Smyk et al, 2019],

the influence of factors on student attrition [Valeeva, Dokuka and Yudkevich, 2007; Gorbunova, 2013; Kochergina and Prakhov, 2016; Osipova, Kolodeznaya and Shevtsov, 2018], factors of attrition of certain groups of students (e.g., engineering training directions [Shmeleva and Frumin, 2020; Maloshonok and Shcheglova, 2020]), the discourse on the reasons for dropout among university teachers is analyzed [Terentyev, Gruzdev and Gorbunova, 2015], predictive models of dropout are proposed [Rusakov, Rusakova and Posokhina, 2018; Zharikov et al. , 2020; Gafarov, Rudneva and Sharifov, 2023], the reasons for the dropout of graduate students are studied [Bekova, 2020], the relationship between the results of the USE and student dropout [Poldin, 2011; Zamkov and Peresetsky, 2013; Havenson and Solovyova, 2014; Timofeeva, 2016].

In general, it is worth to note the absence of any theoretical development of attrition factor models in the domestic scientific literature. Most published papers are empirical and do not rely on a theoretical framework (except for the works [Gorbunova, 2013; Kochergina and Prakhov, 2016; Shmeleva and Frumin, 2020]). These works rely on Tinto's theory of integration, which is dominant in the field of attrition research, however, it has its limitations (the main provisions of this model and its limitations are described in detail in the paper [Gorbunova, 2018]).

The theory of "academic momentum" was chosen to conduct a comparative analysis of student attrition in the United States and Russia for several reasons. It is worth noting that the selected countries represent different educational systems in terms of sanctions for academic failure, the role of extracurricular activities at university, and freedoms in determining one's educational path. Thus, in the American system a student most often makes a decision to leave the university on his/her own, while in Russian universities the vast majority of expulsions are forced [Donets, 2011]. The theory of "academic momentum" focuses attention on the characteristics of the educational process and student's diligence in learning, which are closely related to academic performance. This theory also considers the fact of a student's dropping out as the main dependent variable, rather than his/her decision to leave the university. In addition, this framework can be applied under the conditions of data limitations on universities in Russia (in particular, there are no representative surveys suitable for studying dropout factors, the only data currently available to researchers are administrative data on individual universities, containing information on student enrollment and history of their studies at the university). Finally, the focus of this theoretical framework on the characteristics of the learning process allows the university to concentrate on the factors it can regulate (academic performance, student workload), which is important for the formation of educational policies aimed at student retention.

Research Hypotheses

1. To combine different periodicities in EHA models without dynamic covariates, the use of a technique for reducing the semester system to a trimester system using probability distributions of the occurrence of events will be more effective compared to aggregation to a lower frequency, or interpolation to the highest frequency in terms of graphical display of periodicities, obtaining correct estimates of coefficients for regressors, as well as meaningful interpretation.

2. To combine different periodicities in EHA models with dynamic covariates, the use of stratified proportional odds models (with stratification depending on the type of university) will

be more effective compared to the methods of aggregation to the lowest frequency, interpolation to the highest frequency, and also reduction of the semester system to trimester one using probability distributions of events from the point of view of obtaining correct estimates of coefficients for regressors (both static and dynamic), as well as from the point of view of interpreting model coefficients.

3. Higher “academic momentum” reduces the likelihood of dropping out of a university, both in the USA and in Russia.

4. The impact of “academic momentum” indicators will vary depending on such institutional factors as the level of selectivity of the university, field of study, and the gender of the student.

5. The impact of “academic momentum” indicators will differ between Russia and the United States.

.

Data collection and analysis techniques

Empirical basis of the study

To conduct a comparative analysis of the factors of student attrition from Russian and American universities on the basis of the theoretical framework of "academic momentum", data on the history of students' education in two Russian and eight American universities were used.

The Higher Education Information System database, which aggregates individual students' data for Ohio universities, was used to analyze data on U.S. institutions of higher education[4]. It contains detailed records on socio-demographic characteristics, enrollment characteristics, training courses completion, academic performance, financial aid receipt for each student who was enrolled in Ohio universities. Individual students' information from this database was updated with regional-level data from the Census Bureau's American Community Survey. Ohio universities were selected according to the following characteristics: they are public, have a selective admission system (characterized by competition for entry), are not a branch of an institution, and provide educational services for the mastery of bachelor's degree programs. The analysis was conducted for students enrolled in "full-time" four-year bachelor's degree programs in 2007, which totaled 25339 students.

Data on Russian institutions of higher education were provided by administrative data for two Russian public universities. One of them is a highly selective university located in Moscow and specializes in teaching socio-economic programs. The other institution is medium selective, located in a large regional Russian city, and specializes in engineering and technology. Both universities provided impersonal individual details of students enrolled in full-time bachelor's or specialist programs in 2009. Namely, data were provided on students' socio-economic background (gender, age, region of residence), on their schooling outcomes (type of secondary school, year of graduation, graduation from secondary school, availability of a gold medal for high academic performance), method of admission to higher education institution (by Olympiad, USE, benefits, tuition-paid place), program to which students were admitted in the university, their academic

4 Permission was obtained from the Institutional Review Board of Ohio State University to access data on Ohio's higher education institutions.

performance there and learning trajectories (date and reason for expulsion, academic leave, transfer to another faculty). The total sample amounted to 6553 people.

For comparability purposes, the sampling was limited to "traditional type" students, i.e., first-time, full-time, full-time undergraduate students and students of typical age (16-24 years old in Russia and 18-24 years old in the United States at the time of enrollment). In both countries, the observation period of the student's pathway was limited to two and a half years from the time of their enrollment (two academic years and the fall period of the third year of study)[5].

Methods of data analysis. Principles of analytical model construction

When analyzing data on students' dropping out, we used the methods of event history analysis (EHA), which allow to take into account the temporal aspect of the event occurrence [Allison, 1984; Singer and Willet, 2003], in particularly, censored information and dynamic covariates. The relevancy of their application to this research object is justified by the peculiarities of the data, which, in addition to the fact of student dropout, also contained information about the timing of this event. A group of researchers showed that the use of multiple regression allows to partly solve the problems of censoring and independent dynamic variables, but leads to loss of information and shifts in estimates [Allison, 1984]. Using as a dependent variable an indicator that reflects only the fact of the event occurrence (was expelled - was not expelled) leads to the problem of loss of information about the time of the event occurrence [Hardy and Bryman, 2004]. In particular, we will not know what influences the probability of being dropped in various courses of studies.

In Russian universities, there was an accurate date of a student's expulsion. In US universities, the student's study status was fixed for each period of study - semester or trimester. Discrete EHA models were chosen [DesJardins, Ahlburg and McCall, 1999]. This decision was taken because in American universities the date of withdrawal was fixed in discrete units, while in Russian universities, despite the exact fixation of the day when the order to withdraw the student was issued, the data were tied to the period of the examination session, i.e., in fact, were discrete units. Also, discrete EHA models allow easily to integrate dynamic independent characteristics into the model - in our case, these were records of financial aid receipt by U.S. university students recorded for each period of a student's higher education.

In purpose of constructing the distribution of risk functions of students' dropout in both countries, as well as assessing the differences between risk functions by various parameters (gender, specialty of study, etc.), descriptive EHA methods, specifically, life time tables and the Kaplan-Meier method were used.

To study the linear relationship between “academic momentum” indicators and the probability of student attrition the discrete-time event history model was estimated for each country using a logistic regression model:

[5] The choice of this period was caused by the limited amount of data on Russian universities. According to the results of many researches, it is the first two years of Bachelor's/specialist degree studies that account for the largest percentage of students' attrition (Kolotova, 2011; Donets, 2011).

At the stage of studying the relationship between measures of "academic momentum" and student withdrawal from a university, discrete EHA models of three specifications were used:

1) Model with logit link (further – LL models) [Jenkins, 1995]

$$\frac{h(t_j;x,\beta,\alpha)}{1-h(t_j;x,\beta,\alpha)} = \exp(x'\beta + g(t_j;\alpha)) \quad (1)$$

2) Model with cloglog link (further – CL models) [DesJardins, Ahlburg and McCall, 1999]

$$h(t_j;x,\beta,\alpha) = 1-\exp(-\exp(x'\beta + g(t_j;\alpha))) \quad (2)$$

3) Model of proportional odds (further – PO models) [McCullagh, 1980; Bennett, 1983]

$$\frac{1-S(t;x,\beta,\alpha)}{S(t;x,\beta,\alpha)} = \exp(x'\beta + g(t;\alpha)). \quad (3)$$

where j – observation period,

$h(t_j)$ – risk of event occurrence (conditional possibility of event occurrence in period j, on condition that it has not occurred earlier, and taking into account the values of constant and dynamic covariates (x)),

S (t) – survival function (the probability that an event will occur later than a certain period of time t),

$g(t_j; \alpha)$ – a function reflecting the time dependence (the relationship between the probability of ending a state and the duration t of staying in it) - a vector of parameters of this function)

x – vector of constant and dynamic covariates,

β – vector of coefficients for the corresponding covariates.

Regression models were estimated separately for each country. As a dependent variable in Russian universities on the basis of data on the date of expulsion from the university, a variable was constructed that fixed the fact of a student's expulsion in each quarter of the year of study (1 period was equivalent to 3 months). The U.S. institutions data did not capture the exact date of dropout, but contained information on whether the student was enrolled in a particular academic period at the institution. Based on these records, a variable was constructed to reflect the fact and time of withdrawal (a student was considered to have dropped out if he or she did not attend the institution for a period of at least one year and one academic period).

Thus, the dependent variable in each country was the variable reflecting the fact and time of dropping out of the university, and the independent variables were the indicators of "academic momentum" (see Table 1), as well as control variables. The following indicators of "academic momentum" were considered in our research: the break between graduation and enrollment in a university (the so-called "delayed start"), the student's declaration of a major at the time of enrollment in a university, the presence in the student's curriculum of adaptation courses in the first period (trimester or semester, depending on the university) of the first year of study, academic load in the first academic period of the first year of study - the number of accumulated credits,

academic performance in the first academic period of the first year of study[6], changes in academic performance, and the number of credits accumulated in the first academic period of the first year of study[7][8]. The index "student's identification of his/her major at the moment of enrollment in the university" - was for the first time included in the analysis of the "academic momentum" model.

The following control variables were used in the models for Russian institutions of higher education: student's age at the time of enrollment, student's gender, region of residence before entering the university, staying in a dormitory, group of specialties of study at the time of enrollment, receipt of subsidies in tuition fees from the state due to the results of enrollment in the university, presence of a gold medal for school achievements, university of study. The following control variables were used in the models for U.S. institutions of higher education: student age at the time of enrollment, student gender, student race, student residence in a dormitory during the first academic period, receipt of financial aid during each period of study (grants for low-income students, grants for high academic achievement, and financial aid for working at the university), characteristics of the residence region (proportion of the population with a post-secondary degree, unemployment rate), and the university.

Table 1

Operationalization of indicators in models for Russian and US universities

Indicator	Operationalization of indicators	
	US	Russia
Departure from university – status and time	The input contained data about whether the student studied in a particular semester or trimester at the university. They were transformed into two variables - the fact of departure and the time of departure (semester or trimester of a particular academic year). A student was considered to have dropped out if he did not attend the university for a period of one year and one academic period or more.	At the entrance there was data about the exact date of the student's expulsion. These were converted into two variables – the fact of disposal and the time of disposal (the specific date of disposal was aggregated to the quarter of the year).
Indicators of academic momentum		
Break between leaving school and entering higher education ("delayed start")	A dichotomous variable, which equals to 1 if college enrollment was delayed by at least one year.	A dichotomous variable, which equals to 1 if college enrollment was delayed by at least one year.

6 In Russian higher education institutions the academic performance in the first semester of the first year of study was considered, in US higher education institutions - in the first trimester or the first semester of study, depending on the peculiarities of the academic year system (trimester or semester).

7 In Russian universities, the change in academic performance between the first and second semesters was considered, while in US universities it was between the first and second trimester or the first and second semester, depending on the peculiarities of the academic year system.

8 Not all indicators were included in the model for Russian universities due to the peculiarities of the educational system at the time of the study (the inability to choose a specialty after enrolling in a university, the choice of study load in the first year of education, the lack of a system of remedial courses at the university level).

Continuation of table 1

Reporting major upon university entry	A dichotomous variable that takes the value 1 if the student did not declare a major in the first academic period or chose to study in non-major or liberal arts (liberal arts, general programs).	Wasn't include.
Availability of adaptation courses in the student's curriculum in the first academic period	A dichotomous variable taking the value 1 if the student included adaptation courses in his/her academic plan in the first academic period.	Wasn't include.
Total number of attempted academic credits in the first term	Reflected the total number of attempted credits in the first term and was split into three groups: [12- 15), [15-17), [17, ...).	Wasn't include.
First-term academic performance	Academic performance was expressed using the first-term GPA, which was split into 7 or 10 categories depending on the model specification.	Academic performance was approximated using course failures in the first term, and was split into 4 categories: 0, 1, 2, 3 and more course failures.
Changes in grades over time (between the first and second terms)	The variable reflects whether a student's GPA improved, declined, or remained unchanged between the first and second terms. The changes in grades reflect shifts between the GPA groups between the first and second terms: GPA=0, (0, 1.5), [1.5, 2), [2, 3), [3, 3.5), [3.5, 4].	The variable reflects whether a student's GPA improved, declined, or remained unchanged between the first and second terms. The changes in grades reflect shifts between the average score groups between the first and second terms: [2; 3], (3; 3.5], (3.5; 4], (4; 4.5], (4.5; 5].
Control variables		
Grouped variable of field of study	A set of binary variables.	A set of binary variables.
University	A set of binary variables.	A set of binary variables.
Age	Number of years at time of enrollment.	Number of years at time of enrollment.
Gender	Male or female.	Male or female.
Financing student education	A set of binary variables fixed at every period of study (semester or trimester) with values: receiving a grant based on low family income, a grant for academic achievement, taking out a student loan, covering educational expenses by working for the university.	Binary variable, fixed at moment of enrollement (budget or commerce basis).

Continuation of table 1

Student accommodation on campus	Accommodation in a dormitory or rented housing on the university campus recorded for the first period of study (semester or trimester). A binary covariate.	The interaction of the variables was whether a place was given in a university dormitory in the first academic period (the first quarter of the first year of study) and the region of residence (aggregated categories).
Ethnicity	Student's ethnicity (Caucasian, African American, Hispanic, Asian, other).	Wasn't include.
Characteristics of the student's region of residence before entering the university	Proportion of population with a bachelor's degree or higher, unemployment rate.	Wasn't include.
Scores at school	Wasn't include.	Having a gold medal for successfully completing school.

Both general models for each country and separate models for different types of university selectivity, male and female student gender, and different specialty groups (for each country separately) were constructed, which allow us to conclude on the monitored heterogeneity in these features. In addition, the focus on more homogeneous groups of students improved the validity of the cross-country comparative analysis.

Research findings

Empirical and methodological results were obtained. At first, we describe the results related to the development of the methodology of accounting for different periodicity in EHA on the example of analyzing student dropout.

The development of techniques for matching data with different periodicities

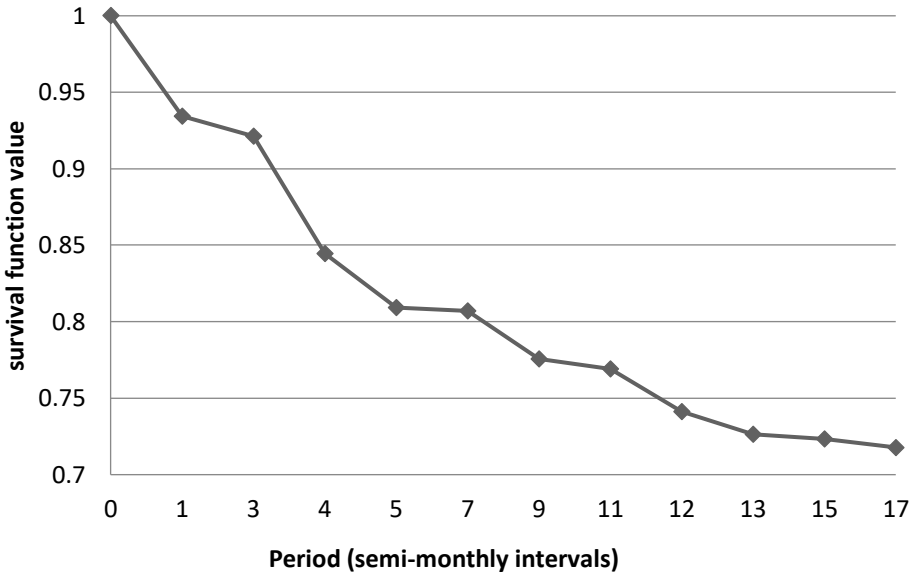
Since *j* - the observation period - was different for U.S. higher education institutions with semester and trimester academic year systems, it became necessary to provide approaches to account for the varying frequency of the boundary control.

We have first suggested techniques that require bringing to a uniform periodicity before building a statistical model, such as: aggregating trimesters and semesters to annual intervals, interpolating trimesters and semesters to semi-monthly intervals, and converting a semester system to a trimester system using distributions of probabilities of occurrence of events.

The technique of aggregating trimester and semester records to annual intervals is the simplest; however, in the case of comparative analysis of attrition data, it leads to a significant loss of information because the observation period is 2.5 years.

In the paper [Kondratjeva, Gorbunova and Hawley, 2017] the application of the technique of interpolation of trimesters and semesters to one and a half month intervals was tested. This

article deals with such a configuration of this technique when an event refers to the first period of the new periodicity under consideration (e.g., if an event occurred in the first semester, it is assigned to the first one and a half month interval of the new periodicity). Thus, this methodology is deterministic, as the decision to which interval of the new periodicity to allocate the occurrence of an event is determined by the researcher. This methodology allows to maintain the detail of the considered features, however, the new periodicity is not naturally used to characterize the process of attrition. In particularly, the interpretation of the survival function graphs is extremely difficult when using this method (pic. 1).



Pic.1. Survival function curve after the procedure of interpolation of semesters and trimesters to one and a half month intervals

In an effort to overcome the mentioned limitations, we first proposed a methodology for converting a semester system to a trimester system using probability distributions of event occurrence. It is described in more detail in the study [Gorbunova and Ulyanov 2016]. The point of this technique is that semester records are reduced to trimester records on the basis of the suggested probability algorithm.

$$p_{11} = \frac{b_1}{b_1 + b_{21}} = \frac{b_1}{a_1}, \quad (4)$$

$$p_{12} = 1 - \frac{b_1}{a_1}, \quad (5)$$

$$p_{23} = \frac{b_3}{b_3 + b_{22}} = \frac{b_3}{a_2}, \quad (6)$$

$$p_{22} = 1 - \frac{b_3}{a_2} \quad (7)$$

where

a_i - the share of dropouts in the i-semester for all universities with semester system of education,

i =1,2 (fall and spring semesters),

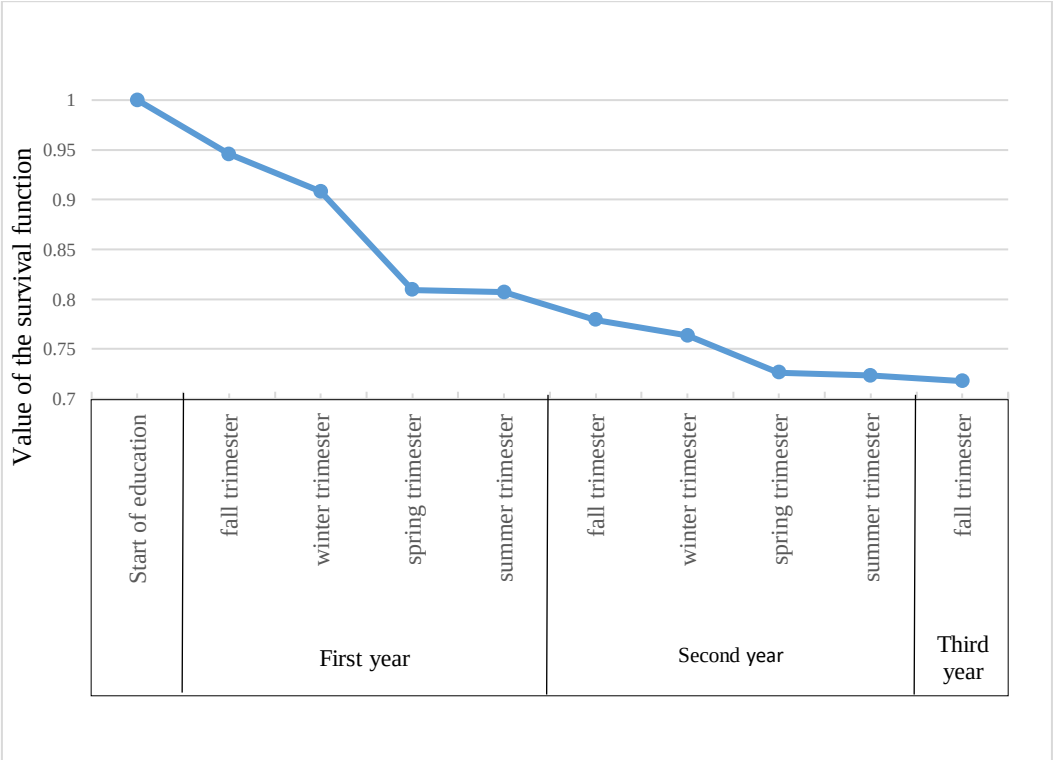
b_j - the share of those who dropped out in the j-trimester for all universities with the trimester system of education,

j=1,2,3 (fall, winter, spring trimesters),

p_{ij} – is the constructed probability with which an event from semester i refers to trimester j.

After calculating the desired probabilities, the process of relating the semester data to the trimester begins. It is performed in the statistical package using the random number function. A new variable is created in the array, which is a random number function (hereinafter referred to as randomvar). Further, a logical condition is set: if the observation is initially referred to the first semester, it will be referred to the first trimester, randomvar =< p11, and to the second trimester if randomvar > p11. The similar procedure is carried out for the second semester.

Approbation of the method of converting the semester system to a trimester system using probability distributions of the occurrence of events on the data on student expulsion showed that it allows to keep the detail of the considered attribute, does not depend on the researcher's decision to which interval of the new periodicity to assign the occurrence of the event, does not create fictitious variables when splitting smaller intervals into larger ones, and offers a periodicity convenient for interpretation and construction of survival rate graphs (all observations are described in trimesters). However, when using this method, there is a problem of violation of causal connections if dynamic covariates are included, as described in the article [Gorbunova and Ulyanov, 2016: 146-148]. According to the results of the conducted validation, this technique is the most suitable for the tested data under the condition of building regression models without dynamic covariates. If the model is built using dynamic covariates, this technique is most suitable at the stage of descriptive data analysis - construction and interpretation of survival tables and graphs of the risk function and survival (pic. 2).



Pic.2. Survival function curve after the semesters-to-trimesters procedure

In investigating techniques that achieve the best results when combining multiple periodicity data, we turned to stratified EHA models, which do not require data to be fit to a single periodicity before estimating a statistical pattern. Strata were identified on the basis of the nature of time dependence (the first stratum included universities where education is divided into semesters, and the second stratum included universities where education is divided into trimesters). Stratification is most often considered in Cox models (see, for example, [Ata and Sözer, 2007]); in our case, the given principle - specification of the function describing the time parameter, g(t;a) separately for semester and trimester universities - was applied to the commonly used logit-linked, cloglog-linked EHA models (hereinafter - LL and CL models), as well as to the not so widely used proportional odds model (hereinafter - PO). It is worth to mention that the stratification principle was first described in application to the three models mentioned above. How these models were defined with stratification is described below:

Model LL:

$$\frac{\mathrm{P}(y_{it}=1)}{\mathrm{P}(y_{it}=0)} = \exp(x_{it}'\beta + z_t^{trim} trim_i'\gamma + z_t^{sem} sem_i'\delta), \quad (8)$$

where

y_{it} — student dropout indicator i while studying t, x_{it} — vector of the explanatory variables, student's and university's characteristics, $trim_i$ — periodicity indicator ($trim_i = 1$, if student i is studying at a university with a trimester system, $trim_i = 0$, If the training is divided into semesters), z_t^{trim} — vector of time period indicators for trimester universities, z_t^{sem} — for "semester-long" universities, β, γ и δ — vectors of estimated coefficients, wherein does not include the free component - it is taken into account separately for different types of universities in the vectors γ и δ.

Model CL:

$$\mathrm{P}(y_{it}=1) = 1 - \exp(-\exp(x_{it}'\beta + z_t^{trim} trim_i'\gamma + z_t^{sem} sem_i'\delta)) \quad (9)$$

The proportional odds model:

$$\mathrm{P}(y_{i1}=1) = \frac{\exp(x_{i1}'\beta + trim_i g_1^{trim} + sem_i g_1^{sem})}{1+\exp(x_{i1}'\beta + trim_i g_1^{trim} + sem_i g_1^{sem})}. \quad (10)$$

The proportional odds model was proposed by D. Cox in 1972 [Cox, 1972] and is well suited for analyzing discrete data. It is also recommended for use when the data contains a large number of “ties” - events occurring at the same time [Cox and Oakes, 1984].

In contrast to the popular proportional hazards model (which is also known as Cox regression), which assumes that the hazard for two individuals at time t is proportional and constant at any point in time (hi(t)/hj(t), where i≠j), in the proportional odds model, risks for different individuals may be disproportionate. The work of [Bennett, 1983] shows that over time, the differences between the risks of events occurring for two individuals decrease.

The results of testing these models are detailed in the study [Gorbunova, Ulyanov and Furmanov, 2017]. It was shown that the proportional odds model has an advantage with regard to the analysis of student dropout both in regard to interpretation and on the basis of the quality of the fitting (the value of the likelihood function) (see Table 2). In particular, the coefficients of the cloglog-model are not interpretable because they consider the process as continuous (while dropouts are related to passing the boundary control and are essentially discrete). The interpretation of logit-model coefficients is related to the education period, i.e. the same coefficient value corresponds to a different in magnitude connection between the odds of attrition and the explanation variable for different education periods. The interpretation of the PO model coefficients is not bound to periodicity because, in contrast to the logit-bound model, it relies on the odds ratios not in a particular period h(t)/(1-h(t)), but on the odds of dropping out up to time t, equal to (1-S(t))/S(t). However, it is impossible to use standard software to estimate the proportional odds model; the researcher needs to carry out the procedure of its estimation independently. The article describes the algorithm of estimation of this model using the STATA program.

Table 2

Estimates of some model coefficients LL, CL и PO

Variable	LL		CL		PO	
	β	exp(*β*)	B	exp(*β*)	β	exp(*β*)
Availability of adaptation courses in the first period of study	0.251	1.285	0.204	1.226	0.278	1.320
Number of credits accumulated during the first study period (basic category - less than 15)						
≥17 credits	–0.268	0.765	–0.239	0.787	–0.299	0.742
[15; 17)	–0.136	0.872	–0.113	0.893	–0.170	0.844
Average scores for the first period of study (6 groups, basic - 4)						
1 group (lowest scores)	2.108	8.231	1.778	5.918	2.825	16.861
2 group	1.319	3.740	1.198	3.313	1.602	4.962
3 group	0.641	1.898	0.596	1.815	0.749	2.115
5 group	–0.484	0.616	–0.470	0.625	–0.539	0.583
6 group (highest scores)	–0.755	0.470	–0.739	0.478	–0.832	0.435
Plausibility logarithm	–22921.524	–22954.799	–22817.686			

Remark. All estimates shown are considerable at the 0.1% level.

Comparing the effectiveness of the developed methods was beyond the scope of this study. However, based on the analysis, we identified several criteria that can guide the researcher when choosing among them:

1) the need to preserve detailed information about the time of the event.

This parameter is not always significant for the researcher, although aggregation of data at lower frequencies allows obtaining less accurate estimates. However, in the case when the number of periods under consideration is quite large (as, for example, in [Chiang, 2012]), the researcher can neglect the accuracy and reduce the data to a single periodicity (lower), and then conduct

descriptive and regression analysis using discrete semiparametric EHA models with logit link [Allison 1984; Singer and Willet 2003].

If it is important to preserve detailed information about the time of occurrence of an event, then it is necessary to choose among the other five methods. The choice between them will depend on the criteria listed below.

2) The importance of preserving the natural periodicity of data in descriptive statistics.

Only two of the methods we describe allow us to obtain distributions of risk and survival functions in a periodicity that is convenient for interpretation and natural for the phenomenon under consideration - aggregation to annual intervals and reduction of the semester system to a trimester system using probability distributions of the occurrence of events.

In particular, the interpretation of graphs at a one and a half month periodicity (in the case of interpolation to data of the highest frequency - in our case, one and a half month intervals) is quite difficult; a month and a half periodicity is not natural for describing learning processes at a university. Stratification methods also do not solve the problem of analytical presentation of data in a periodicity convenient for interpretation; data are presented separately at different periodicities.

Thus, if it is important to preserve the detail of the time of occurrence of an event, the most appropriate would be to use the technique of reducing the semester system to the trimester system using probability distributions for the occurrence of events. If the loss of information is not significant, aggregation to the lowest frequency is suitable for solving this problem.

3) The need to include dynamic covariates in the regression model and obtain correct estimates from them.

The results of our analysis showed that the best solution for implementing discrete EHA models with dynamic covariates would be to use a proportional odds model with stratification by objects with different frequencies (in our case, universities with semester and trimester education systems). Firstly, the interpretation of the coefficients of this model is not tied to periodicity, unlike with the logit link, but is based on odds ratios up to time t, equal to (1-S(t))/S(t)). Secondly, from the point of view of the criteria for the quality of model fit (the value of the likelihood function), it gives the best results (Table 2). However, in a situation where the probabilities of an event occurring are low (as in the considered example with the expulsion of students), and obtaining interpretable coefficients is not important, the LL model can also be applied (the differences in the coefficients for regressors in the semester and trimester systems will be small).

Results on factors of departure from Russian and US universities

Now let us describe the results related to the empirical task of revealing the factors of student attrition in Russian and American universities. The hypothesis of the study was confirmed, indicating that higher "academic momentum" reduces the probability of student attrition both in Russia and in the USA. Detailed results are presented in the article [Kondratjeva, Gorbunova and Hawley, 2017]. Specifically, in the U.S. sample, which includes all eight institutions, a "delayed start" (a break of a year or more between high school graduation and the start of higher education), the inclusion of adaptation courses in the individual education plan, and a student's postponement of declaring his or her major increase the probability of dropping out. While higher academic load

in the first period of study, higher grade point average in the first period of study, as well as improved academic performance between the first and second periods of study reduce the probability of dropping out. According to the sample of two Russian universities, a "delayed start" and the presence of failed courses in the first semester of study increase the probability of expulsion, while the increase in the average grade point between the first and second semester reduces it. Statistical tables with the results of the analysis are presented in Appendix 2.

It is worth to mention that a higher probability of students with a "delayed start" dropping out of university is also reported in the paper [Attewell, Heil and Reisel, 2012]. However, this conclusion is not confirmed in the work [Martin et al, 2013] - in this study on the data on students of Australian universities it was shown that "delayed start" is positively associated with the probability of graduation. It should be noted that "delayed start" may be a consequence of other variables, and when interpreting it is worth considering the educational experience of the student, with which he came to higher education, as well as the reason why he takes this break. This aspect is studied in detail in the study [Roksa and Velez, 2012].

The results demonstrating that a higher academic workload is associated with a lower risk of dropping out are supported by the works [Adelman, 1999; 2006]. However, these findings are not supported by the paper [Attewell, Heil and Reisel, 2012].

Our results regarding the inclusion of adaptation courses in the curriculum in the first year of study, as well as the negative relationship of higher academic performance and attrition are in accordance with the findings obtained in the works [Adelman, 1999; 2006].

Our analysis shows the heterogeneity of the link between "academic momentum" and student attrition depending on the level of selectivity of the university, specialty of study and student gender. Detailed results are presented in the article [Kondratjeva, Gorbunova and Hawley, 2017] and are summarized in Appendix 1.

There were also recorded differences between the two countries in the effect of academic momentum on student attrition in the examined subsamples, in particular, with regard to the indicator "break between graduation and enrollment in higher education". In particular, this indicator does not have a meaningful effect on attrition for technical and natural science students in US institutions[9], but it is significant for these specialties in Russian universities.

Key results of the study (positions put forward for defense)

1. The paper is the first to introduce a methodology for reducing *the semester system of education to a trimester system using probability distributions of event occurrence.* Testing of this technique with data on student departures showed that its main advantage, compared to the method of data interpolation to the highest frequency, is the probabilistic approach to assigning a new periodicity to data of lower frequency, as well as the presentation of data in a natural periodicity for the object, which is important when interpreting life tables and graphically

[9] Though, this conclusion is inconsistent with the results of a more recent study of technical and natural-scientific students, which established that a higher academic workload at the beginning of university studies is positively related to the likelihood of completing a program [Zhang, 2019].

displaying risk and survival functions. A limitation of this technique is the inclusion of dynamic covariates in the model.

2. Applied to the problem of building a regression model with dynamic covariates, the application of a *stratified proportional odds model* to duration records with different periodicity is the best solution in view of the interpretation of the model coefficients and the quality criteria of the model fit (the value of the plausibility function). The limitation of this approach is the inability to use standard software for its assessment; the researcher must implement the assessment procedure independently.

3. We proposed the following criteria for choosing one of the six considered techniques for combining data with different periodicities in EHA: the need to preserve detailed information about the time of the event; the importance of preserving the natural periodicity of data in descriptive statistics; the need to include dynamic covariates in the regression model and obtain correct estimates from them.

4. The theory of "academic momentum" appeared to be applicable for the comparative analysis of student attrition in Russia and the USA, because the theoretical provisions that higher "academic momentum" reduces the risk of dropping out of universities are confirmed in empirical researches conducted in the two countries. Specifically, higher "academic momentum" reduces the probability of students' dropping out of universities both in the USA and in Russia. Precisely, in the analyzed Russian universities, the presence of a break between graduation from school and enrollment in university, and the presence of failed subjects in the first semester of study increase the probability of expulsion (wherein, there is an increase in the probability of expulsion with the increase in the number of failed subjects), while the increase in the average score between the first and second semester of study reduces it. In the US institutions in Ohio, a "delayed start" (a break of a year or more between high school graduation and the start of higher education), the inclusion of adaptation courses in the individual study plan, and a student's postponement of declaring his or her major increase the probability of dropping out of higher education. While higher academic load in the first period of study, higher grade point average in the first period of study, as well as improved academic performance between the first and second periods of study decrease the probability of dropping out.

5. The relationship between “academic impulse” and student attrition varies depending on the level of selectivity of the university, the specialty of study and the gender of the student.

6. Differences were recorded between the two countries in the effect of academic momentum on student attrition relative to the delayed start indicator. This indicator does not have a significant impact on the departure of students in technical and natural sciences at US universities, but is significant for these specialties at Russian universities.

Limitations of the work

This study has its limitations in conducting an empirical comparison of student attrition between Russian and American universities.

Firstly, the empirical results of the research cannot be expanded to all Russian and American universities. In the case of Russia, the data are from two different types of higher

education institutions, and the results may be representative of students enrolled in full-time bachelor's degree programs only in these institutions. In the case of U.S. institutions, the data are from eight public institutions in Ohio - and can be applied to the population of students enrolled in full-time undergraduate programs of that state, or another U.S. state with similar educational system characteristics. To obtain more reliable and representative results, greater cooperation among U.S. states (many of which have a centralized system for registering student data) is required, as well as the creation of a centralized system for collecting data on students at Russian universities.

Second, the study does not take into account some unobserved characteristics such as student employment, marital status, academic motivation, etc. Although the administrative data contained accurate information about a student's marks and academic statuses and provided information on the entire population of students, many of the observed characteristics that could have influenced student attrition were not contained in the records and therefore were not included in the model.

Third, the observation period is limited to two and a half years of study, so the results cannot be extended to the attrition of students in their third and subsequent years of study. Also, since dropouts are examined over a 2.5-year period, information on whether these students returned to study beyond the period under consideration is left out.

Fourth, it is worth taking into account that student performance varies across different universities and majors of study, which is only partially accounted for in our models by control variables (university, group of majors of education).

Fifth, it is worth noting the limitations of conducting cross-country comparisons. In particular, cohorts of different years of admission were taken (2007 in the USA and 2009 in Russia), which was due to the availability of data. In addition, the lower age limit for students differed (16 years in Russia and 18 years in the USA), which is due to the characteristics of the educational systems of the two countries. Also, the set of covariates in the models for the two countries differed, which was also due to their availability. Finally, the model included different numbers of “academic momentum” indicators. In the case of the Russian model, three indicators were included, while in the model for US universities there were six indicators. However, given these limitations, this paper is the first to present the application of academic momentum theory in a comparative context between two countries.

Finally, stability and the effectiveness of the proposed methods was beyond the focus of this study, however, it may be the subject of further investigation.

Conclusion

This study was devoted to the development of techniques for combining data with different periodicities within the framework of event occurrence analysis. The proposed techniques were tested on the data of a comparative research of student attrition factors in Russian and American universities, in which it was necessary to combine trimester and semester records.

Two approaches were proposed within the framework of this task. One approach allowed to bring the data to a single periodicity before constructing the statistical model, while the second approach allowed to solve the problem of bringing together different periodicities simultaneously

with the estimation of the statistical model. In each approach, three data matching techniques were considered. In the first approach, a new technique was proposed to convert a semester-based learning system to a trimester-based system using probability distributions of event occurrence. It was shown that it allows to preserve the detail of the considered feature - the time of the event occurrence, and it is also well suited to describe the survival function on the combined sample with different periodicities. However, its application causes difficulties when a regression model with dynamic covariates is considered. To overcome this limitation, stratified EHA models were considered in which different periodicities represented separate strata. Stratification was considered earlier in Cox's models, but for the first time it was considered in models with logit-link, cloglog-link and proportional odds models. The paper examined the particularities of the interpretation of these models as applied to the analysis of student attrition factors and compared their effectiveness in terms of the quality of fit. It was concluded that the use of stratified proportional odds model is the most appropriate for analyzing student attrition factors. Testing these approaches on student attrition data resulted in conclusions about what problems the proposed approaches can and cannot solve, as well as recommendations on how to choose among them. Although the proposed approaches were formulated for combining different periodicities on U.S. data, they can be successfully applied to solve a similar problem in other educational systems as well as combining data with different periodicities for analyzing other social objects.

In addition to obtaining methodological conclusions, results were obtained concerning the applicability of the "academic momentum" model used for the first time in the comparative study and on Russian data. In particular, the main statements of this theory were confirmed on the basis of our study evidence. Namely, the higher speed with which a student starts mastering a program in higher education reduces the probability of his/her dropping out of the university. This conclusion was based on the analysis of such indicators of academic momentum as the break between graduation and enrollment in the university, the student's announcement of a major at the time of enrollment in the university, the presence in the student's curriculum of adaptation courses in the first period (trimester or semester, depending on the university) of the first year of study, the workload in the first academic period of the first year of study - the number of credits accumulated, academic performance in the first academic period of the first year of study, and changes in academic performance between the first and second academic periods of study.

Additionally, the results were obtained on how the influence of the indicators of "academic momentum" differs depending on such institutional factors as the level of selectivity of the university, specialty of study, as well as the gender of the student. Interdisciplinary differences are primarily related to the peculiarities of educational programs, in particular, the complexity of the disciplines studied, as well as the requirements for the preparation of university applicants. Differences connected with the selectivity of higher education institution can be explained by the level of preparation of students entering undergraduate programs, demands imposed on students during their studies in the university, peculiarities of formal and informal rules of expulsion. Gender differences in the effects of "academic momentum" can be related to the different roles of men and women in providing financial well-being of the family, their diligence in studying, choosing certain specialties of training.

There were also recorded differences between the two countries in the effect of academic momentum on student attrition in the subsamples under consideration, in particular, with regard to the indicator "break between graduation and enrollment in higher education". The observed differences between the countries can be explained by the peculiarities of the educational systems of the two countries under study, for example, greater opportunities for academic assistance in US universities for students with learning difficulties, as well as cultural peculiarities (lower distance between teacher and student).

The results of the analysis have shown that the model of "academic momentum" can be successfully used in the study of student attrition factors both in the Russian and international context. Its application will be more relevant as the higher education system (including the Russian one) becomes more globalized, and as the educational process becomes more flexible: providing students with the opportunity to postpone the choice of a specialty of study, to choose the study load in each academic period. Many of these changes are already noted today in Russia, such as the switch to a two-tier system, the European system of crediting academic units, the establishment of individual curricula that include both basic disciplines and elective courses, the possibility of including mass open online courses in the curriculum, the introduction of adaptation courses by individual universities, etc. [Resolution..., 2017].

Thus, new methodological and empirical results were obtained in our work. The results on the development of methods for combining different periodicities in EHA can be used both in researches on students' dropout factors and in other fields of study that use longitudinal data with different periodicities. The empirical outcomes can be used both for the formation of educational policies aimed at reducing the proportion of students who do not complete higher education, the adaptation of new educational policies, and the evolution of the theory of students' attrition in the international context.

Bibliography

1) Adelman, C. (1999) *Answers in the Tool Box: Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment*. Washington, DC: U.S. Department of Education.

2) Adelman, C. (2006) *The toolbox Revisited: Paths to Degree Completion from High School through College.* Washington, DC: U.S. Department of Education.

3) Allison, P. D. (1984) *Event history analysis*. Sage Publications.

4) Artamonova, A.V. and Mitrofanova, E.S. (2016) Cohabitation in Russia: "trial" union or an independent social institution. *Monitoring of Public Opinion: Economic and Social Changes*, 1, p. 126—145.

   https://doi.org/10.14515/monitoring.2016.1.04

5) Ata N., Sözer M. T. Cox regression models with nonproportional hazards applied to lung-cancer survival data // Hacettepe Journal of Mathematics and Statistics. 2007. Vol. 36. № 2. P. 157–167.

6) Attewell, P., Heil, S. and Reisel, L. (2012) What is academic momentum? And does it matter? *Educational Evaluation and Policy Analysis*, 34(1), p. 27-44.

   https://doi.org/10.3102/0162373711421958

7) Attewell, P. and Jang, S. H. (2013) Summer Coursework and Completing College. *Research in Higher Education Journal*, 20, pp. 117-141.

8) Bandura, A. (1997) *Self-efficacy: The exercise of control*. New York: Freeman.

9) Bandura, A. (2001) Social cognitive theory: An agentive perspective. *Annual Review of Psychology*, vol. 52, pp. 1–26.

   https://doi.org/10.1146/annurev.psych.52.1.1

10) Bean, J. and Eaton, S. B. (2001) The Psychology Underlying Successful Retention Practices. *Journal of College Student Retention: Research, Theory and Practice*, 3 (1), pp. 73–89.

    https://doi.org/10.2190/6R55-4B30-28XG-L8U0

11) Becker, G. S. (1964) *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. New York: National Bureau of Economic Research.

12) Bekova, S. (2020) Academic Suicide: Scenarios of Doctoral Student Attrition in Russia. *Educational Studies Moscow*, 2, pp. 83-109.

    https://doi.org/10.17323/1814-9545-2020-2-83-109

13) Bennett, S. (1983) Log-logistic regression models for survival data. Journal of the Royal Statistical Society. Series C, 32 (2), pp. 165–171.

    https://doi.org/10.2307/2347295

14) Berger, J. B. (2000) 'Optimizing Capital, Social Reproduction, and Undergraduate

Persistence: A Sociological Perspective'. In Braxton, J. M. (ed.), *Reworking the Student Departure Puzzle*. Nashville, TN: Vanderbilt University, pp. 95– 124.

15) Braxton, J. M. (2000) (ed.) *Reworking the Student Departure Puzzle. Nashville*. TN: Vanderbilt University.

16) Braxton, J. M., Doyle W. R., Hartley III H. V., Hirschy A. S., Jones W. A. and McLendon M. K. (2013) *Rethinking college student retention*. John Wiley & Sons.

https://doi.org/10.1353/csd.2014.0061

17) Chiang, S.-C. (2012) *Applying Event History Analysis to Investigate the Impacts of Developmental Education on Emerging Adults' Degree Completion*. Ph.D. dissertation, Ohio State University. Available at: https://etd.ohiolink.edu/acprod/odb_etd/ws/send_file/send?accession=osu1331061887&disposition=inline [Accessed 20 June 2024]

18) Cox, D. R. (1972) Regression models and life-tables. *Journal of the Royal Statistical Society, Series B*, 34, pp. 187–220.

https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

19) Cox, D. R. and Oakes D. (1984) *Analysis of Survival Data*. London: Chapman and Hall.

https://doi.org/10.1002/bimj.4710290119

20) Davidson, J. C. and Blankenship, P. (2017) Initial academic momentum and student success: Comparing 4- and 2-year students. *Community College Journal of Research and Practice*, 41(8), pp. 467-480.

https://doi.org/10.1177/1521025119826245

21) Denisova, I. A. (2010) Adult Mortality in Russia: A Microanalysis. *Economics of Transition*, 2(18), pp. 333-363.

22) Denisova, I. A. (2017) Exit of older age cohorts from the labor market: analysis by survival methods. *Population and Economics*, 18(2), pp. 22-49.

23) Deci, E. L. and Ryan, M. R. (1991) 'A Motivational Approach to Self: Integration in Personality'. In Dienstbier R. *Nebraska Symposium on Motivation*, 38. Lincoln, NE: University of Nebraska, pp. 237–288.

24) Demetriou, C. and Schmitz-Sciborski, A. (2011) Integration, Motivation, Strengths and Optimism: Retention Theories Past, Present and Future. In: *Proceedings of the 7th National Symposium on Student Retention*. Norman, OK: University of Oklahoma, pp. 300–312

Available at: https://www.academia.edu/34618701/Integration_Motivation_Strengths_and_Optimism_Retention_Theories_Past_Present_and_Future [Accessed 20 June 2024].

25) DesJardins, S. L., Ahlburg, D. A. and McCall, B. P. (1999) An event history model of

student departure. *Economics of Education Review*, 18(3), pp. 375-390.

https://doi.org/10.1016/S0272-7757(98)00049-1

26) Donec, E. (2011) Opyt issledovaniya studencheskiy otchisleniy na primere MGU [Research experience of student deductions on the example of Moscow State University]. *Monitoring universiteta [University monitoring]*, 6, pp. 33–38.

27) Douglas, D. and Attewell, P. (2014) The bridge and the troll underneath: Summer bridge programs and degree completion. *American Journal of Education*, 121(1), pp. 87-109.

https://doi.org/10.1086/677959

28) Evans, M.D.D. (2005) Where are we now? Real-time estimates of the macroeconomy. *International Journal of Central Banking*, 1(6), pp. 127–175.

https://doi.org/10.2139/ssrn.646103

29) Foroni, C. and Marcellino, M.G. (2013) *A Survey of Econometric Methods for Mixed-Frequency Data*. Norges Bank (Research Working Paper 0). Available at: https://www.norges-bank.no/contentassets/5f6ee92052ad49918658826720a8bed9/norges_bank_working_paper_2013_06.pdf?v=09032017123143 [Accessed 20 June 2024]

30) Gafarov, F., Rudneva, Y. A. and Sharifov, U. (2023) Prognosticheskoye modelirovaniye v vysshem obrazovanii: opredeleniye faktorov akademicheskoy uspevayemosti [Predictive Modeling in Higher Education: Identifying Factors in Academic Performance]. *Vyssheye obrazovaniye v Rossii*, 32(1), pp. 51-70.

https://doi.org/10.31992/0869-3617-2023-32-1-51-70

31) Ghysels, E., Santa-Clara, P. and Valkanov, R. (2006) Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometric*, 131 (1), pp. 59–95.

https://doi.org/10.1016/j.jeconom.2005.01.004

32) Ghysels, E. (2016) Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics*, 193 (2), pp. 294–314.

https://doi.org/10.1016/j.jeconom.2016.04.008

33) Gorbunova, E.V. (2013) Vliyaniye adaptatsii pervokursnikov k universitetu na veroyatnost' ikh otchisleniya iz vuza [Impact of adaptation of freshmen to the university on the likelihood of their expulsion from the university]. *Universitas*, 2, pp. 59–84.

34) Gorbunova, E.V. and Kondratjeva, O.S. (2013) Analiz gendernykh razlichiy v vybytii iz vuza rossiyskikh i amerikanskikh studentov programm bakalavriata [Analysis of gender differences in the dropout of Russian and American undergraduate students]. *Universitas*, 3, pp. 48-69.

35) Gorbunova, E.V. and Ulyanov, V.V. (2016) Diskretnyye modeli analiza nastupleniya sobytiy: razrabotka podkhodov k sovmeshcheniyu dannykh, imeyushchikh raznuyu periodichnost' [Discrete models of Event History Analysis: elaboration of approaches to combining data with different periodicities]. Sotsiologiya: metodologiya, metody, matematicheskoye modelirovaniye [Sociology: Methodology, Methods, Mathematic Modelling], 43, pp. 128-153.

36) Gorbunova, E.V., Ulyanov, V.V. and Furmanov, K. K. (2017) Using data from universities with different structure of academic year to model student attrition. *Applied Econometrics*, 45, pp. 116-135.

37) Gorbunova, E.V. (2018) Student attrition from universities: research in Russia and the United States of America. *Educational Studies*, 1, pp.110-131.

https://doi.org/10.17323/1814-9545-2018-1-110-131

38) Gruzdev, I.A, Gorbunova, E.V. and Froumin, I.D. (2013) Studencheskiy otsev v rossiyskikh vuzakh: k postanovke problemy [Student Dropout in Russian Universities: Towards a Problem Statement]. *Educational Studies*, 2, pp. 67-81.

https://doi.org/10.17323/1814-9545-2013-2-67-81

39) Habley, W. R., Bloom, J. L. and Robbins S. (2012) *Increasing persistence*. Somerset: John Wiley & Sons.

40) Hardy, M. and Bryman, A. (2004) *Handbook of data analysis*. London: Sage Publications.

41) Jenkins, S. (1995) Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics*, 57, pp. 129–138.

https://doi.org/10.1111/j.1468-0084.1995.tb00031.x

42) Kehm, B. M., Larsen, M. R. and Sommersel, H. B. (2019) Student dropout from universities in Europe: A review of empirical literature. *Hungarian Educational Research Journal*, 9(2), pp. 147-164.

https://doi.org/10.1556/063.9.2019.1.18

43) Khavenson, T.E. and Solovyova, A.A. (2014) Relationship between the results of the Unified State Exam and academic performance at a university. *Educational Studies Moscow*, 1, pp. 176-199.

https://doi.org/10.17323/1814-9545-2014-1-176-199

44) Kim, J.S. (2003) Maximum Likelihood Estimation for the Proportional Hazards Model with Partly Interval-censored Data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology*), 65, 2, pp. 489–502.

https://doi.org/10.1111/1467-9868.00398

45) Kochergina, E. and Prakhov, I. (2016) Relationships between Risk Attitude, Academic Performance, and the Likelihood of Drop-outs. *Educational Studies Moscow*, pp. 206–

228.

https://doi.org/10.17323/1814-9545-2016-4-206-228

46) Kolotova, E. V. (2011) Izucheniye otchisleniy sredi studentov bakalavriata/spetsialiteta NIU VSHE [Study of dropouts among undergraduate / specialty students at the Higher School of Economics] In: *Sotsiologicheskiye metody v sovremennoy issledovatel'skoy praktike* [Sociological methods in modern research practice]. M.: NRU HSE. P. 271–279.

47) Kondratjeva, O., Gorbunova, E. and Hawley, J. D. (2017) Academic Momentum and Undergraduate Student Attrition: Comparative Analysis in US and Russian Universities. *Comparative Education Review*, 3, pp. 607-633.

https://doi.org/10.1086/692608

48) Kosmützky, A. and Nokkala, T. (2014) Challenges and trends in comparative higher education: an editorial. *Higher Education*, 67(4), pp. 369-380.

https://doi.org/10.1007/s10734-013-9693-x

49) Kuh, G.D. and Love, P.G. (2000) 'A Cultural Perspective on Student Departure'. In Braxton, J.M. (ed.), *Reworking the Student Departure Puzzle*. Nashville, TN: Vanderbilt University, pp. 196–212.

50) Kuh, G. W., Kinzie, J., Schuh, J. and Whitt, E. (2010) *Student success in college: Creating conditions that matter (2nd ed.)* San Francisco, CA: Jossey-Bass.

51) Kuzminov, Y. and Yudkevich, M. (2007) Universitety v Rossii i v Amerike: razlichiya akademicheskikh konventsiy [Universities in Russia and America: Differences between Academic Conventions]. *Voprosy ekonomiki [Economic issues]*, 4, pp. 141-158.

52) Law, C.G. and Brookmeyer, R. (1992) Effects of Mid-point Imputation on the Analysis of Doubly Censored Data. *Statistics in Medicine*, 11(12), pp. 1569–1578.

https://doi.org/10.1002/sim.4780111204

53) Maloshonok, N. G. and Shcheglova, I. A. (2020) The role of gender stereotypes in the dropout of engineering and technical students. *Monitoring of public opinion: economic and social changes*, 2, pp. 273-292.

https://doi.org/10.14515/monitoring.2020.2.945

54) Manyanga, F., Sithole, A. and Hanson, S. M. (2017) Comparison of Student Retention Models in Undergraduate Education from the Past Eight Decades. *Journal of Applied Learning in Higher Education*, 7, pp. 30-42.

https://doi.org/10.57186/jalhe_2017_v7a3p30-39

55) Marginson, S. and Rhoades, G. (2002) Beyond national states, markets, and systems of higher education: A glonacal agency heuristic. *Higher Education*, 43(3), pp. 281– 309.

https://doi.org/10.1023/A:1014699605875

56) Martin, A. J., Wilson, R., Liem, G. A. D. and Ginns, P. (2013) Academic momentum at university/college: Exploring the roles of prior learning, life experience, and ongoing performance in academic achievement across time. *The Journal of Higher Education*, 84(5), pp. 640-674.

https://doi.org/10.1353/jhe.2013.0029

57) McCullagh, P. Regression models for ordinal data (1980). *Journal of the Royal Statistical Society Series B*,42 (2), pp. 109–142.

https://doi.org/10.1111/j.2517-6161.1980.tb01109.x

58) Mitrofanova, E. S. (2020) (Not) time to grow up: how the age of onset of debut biographical events changes among Russians. *Demographic Review*, 7(4), pp. 36-61.

https://doi.org/10.17323/demreview.v7i4.12043

59) Organization for Economic Cooperation and Development (OECD) (2010) *Education at a Glance 2010: OECD Indicators*. OECD Publishing. Available at: www.oecd.org/edu/eag2011 [Accessed 20 June 2024].

60) Osipova, N. G., Kolodeznaya, G. V. and Shevtsov, A. N. (2018) On the patterns and reasons for expulsions at the university and the motivation of students' educational activities. *Education and Science*, 6, pp. 158-182.

https://doi.org/10.17853/1994-5639-2018-6-158-182

61) Pascarella, E. T. and Terenzini, P. T. (1991) *How College Affects Students: Findings and Insights From Twenty-Years of Research*. San Francisco: Jossey-Bass Publishers.

62) Peresetsky, A. and Davtyan, M. (2011) Efficiency of the Unified State Exam and Olympiads as a tool for selecting applicants. *Applied Econometrics*, 3(23), pp. 41-56.

63) Poldin, O. V. (2011) Forecasting academic performance at a university based on the results of the Unified State Exam. *Applied econometrics*, 1, pp. 56-69.

64) *Postanovleniye ot 26 dekabrya 2017 g. № 1642 Ob utverzhdenii gosudarstvennoy programmy Rossiyskoy Federatsii "Razvitiye obrazovaniya" [Resolution of December 26, 2017 No. 1642 On approval of the state program of the Russian Federation "Development of education"]* (2017) Moscow. Available at: http://static.government.ru/media/files/313b7NaNS3VbcW7qWYslEDbPCuKi6lC6.pdf [Accessed 20 June 2024].

65) Roksa, J. and Velez, M. (2012) A late start: Delayed entry, life course transitions and bachelor's degree completion. *Social forces*, 90(3), pp. 769-794.

https://doi.org/10.1093/sf/sor018

66) Seidman, A. (Ed.) (2005) *College student retention*. Westport, CT: Praeger.

67) Rusakov, S.V., Rusakova, O.L. and Posokhina, K.A. (2018) Neural network model for predicting the risk group according to the academic performance of first-year students.

Modern information technologies and IT education, 14(4). pp. 815-822.

https://doi.org/10.25559/SITITO.14.201804.815-822

68) Rutkevich, M. (2002). *Sotsiologiya obrazovaniya i molodezhi: Izbrannoye (1965–2002). [Sociology of Education and Youth: Selected Works*] M.: Gardariki, 2002.

69) Saponova, A.V. and Kulikov, S.P. (2021) Integration of survey data and digital traces: a review of the main methodological approaches. *Sociology: methodology, methods, mathematical modeling*, 53, pp. 117-164.

https://doi.org/10.19181/4m.2021.53.4

70) Shmeleva, E.D. and Frumin, I.D. (2020) Factors of dropout of engineering and technical students in Russian universities. *Educational Studies Moscow*, 3, pp. 110-136.

https://doi.org/10.17323/1814-9545-2020-3-110-136

71) Singer, J. D. and Willett, J. B. (2003) *Applied longitudinal analysis: modeling change and event occurrence*. Oxford university press.

72) Sinyavskaya, O.V., Chervyakova, A.A. and Gorvat, E.S. (2022) Analysis of factors of exit from the labor market at the age of 45 years and older in Russia: the role of employment characteristics, job changes and education. *Sociological Journal*, 28(2), pp. 50-72.

https://doi.org/10.19181/socjour.2022.28.2.8986

73) Smyk, A.F., Prusova, V.I., Zimanov, L.L. and Solntsev, A.A. (2019) Analysis of the scale and reasons for student dropout at a technical university. *Higher education in Russia*, 6, pp. 52-62.

https://doi.org/10.31992/0869-3617-2019-28-6-52-62

74) Terentyev, E., Gruzdev, I. and Gorbunova, E. (2015) The trial is coming: teachers' discourse on student dropout. *Educational Studies Moscow*, 2, pp. 129-151.

https://doi.org/10.17323/1814-9545-2015-2-129-151

75) Timofeeva, A.Yu. and Avrunev, O.E. (2016) Student expulsions during the learning process: the explanatory power of the Unified State Exam. In: *We continue the traditions of Russian statistics. Collection of reports of the I Open Russian Statistical Congress*. Novosibirsk, 20–22 October 2015. Novosibirsk: Novosibirskiy gosudarstvennyy universitet ekonomiki i upravleniya «NINKH», pp. 133–140.

Available at:: https://nsuem.ru/library/resources/sborniki-s-mezhdunarodnym-uchastiem/Мы%20продолжаем%20традиции%20российской%20статистики_Сборник%20докладов_ОРСК%20т_4.pdf [Accessed 20 June 2024].

76) Tinto, V. (1975) Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, 45(1), pp. 89-125.

77) Tinto, V. (1987) *Leaving college: Rethinking the causes and cures of student attrition*.

Chicago: University of Chicago Press.

78) Tyndik, A. and Biryukova, S. (2015) Prevalence and determinants of childlessness in Russia and Moscow. *Genus*, 71(1), pp. 1-22.

https://doi.org/10.4402/genus-605

79) Valeeva, D., Dokuka, S. and Yudkevich, M. (2013) How Academic Failures Break Up Friendship Ties: Social Networks and Retakes. *Educational Studies Moscow*, 1, pp. 8–24.

80) Wang, X. (2015) Pathway to a baccalaureate in STEM fields: Are community colleges a viable route and does early STEM momentum matter? *Educational Evaluation and Policy Analysis*, 37(3), pp. 376-393.

https://doi.org/10.3102/0162373714552561

81) Wang, X., Chan, H. Y., Phelps, L. A. and Washbon, J. I. (2015) Fuel for success: Academic momentum as a mediator between dual enrollment and educational outcomes of two-year technical college students. *Community College Review*, 43(2), pp. 165-190.

https://doi.org/10.1177/0091552115569846

82) Zamkov, O.O. and Peresetsky, A.A. (2013) Unified State Examination and academic success of undergraduate students at ICEF National Research University Higher School of Economics. *Applied econometrics*, 30 (2), pp. 93-114.

83) Zhang, Y. L. (2019) Early Academic Momentum: Factors Contributing to Community College Transfer Students' STEM Degree Attainment. *Journal of College Student Retention: Research, Theory & Practice*, pp. 1–30.

https://doi.org/10.1177/1521025119881130

84) Zharikov, A.V., Zhuravlev, E.V., Zhurenkov, O.V., Kozlov, D.Yu., Maslennikova, I.S., Savchenko, V.G., Fomenko, A.P. and Chupina, T.V. (2020) Computer model for predicting student expulsion based on admissions campaign data. *High-performance computing systems and technologies*, 4(23), pp. 126-131.

85) Zimmerman, B. J. (2000) Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25, pp. 82–91.

https://doi.org/10.1006/ceps.1999.1016

Appendix 1

Conclusions on the influence of indicators of "academic momentum" on student attrition in the context of the countries studied, field of study, gender, selectivity of universities

A. "Delayed start".

One of the results concerns the "delayed start" indicator. Firstly, in the USA, the greatest effect of this indicator is observed in highly selective universities, while in medium-selective universities this indicator is insignificant, and in low-selective universities the coefficient is significant, but its effect is lower. In Russian universities, this indicator is significant in universities of both types of selectivity, however, in a highly selective university its effect is lower compared to a medium-selective university.

Interestingly, the "delayed start" effect also does not appear for females in US universities, however, it is significant for males.

The "delayed start" effect is also not significant for students of technical and natural science fields of study in US universities, but is significant for these fields in Russian universities.

B. Performance indicators, change in performance.

Another result concerns the indicator "academic performance in the first period of study." In universities in both countries, this indicator is one of the strongest predictors of student departure and is significant in all subsamples studied (depending on the selectivity of the university, the student's gender, and field of study). Differences are observed only in the severity of the effect of a specific level of performance. Thus, in American highly selective universities there is a higher probability of expulsion for students with very low academic performance (the average score for the first academic period is from 0 to 1.5 points, according to a 4-point system), and a lower probability of expulsion for students with the highest level academic performance (from 3.5 to 4 points), compared with universities of medium and low selectivity. In Russian universities, the results of a descriptive analysis showed that in a highly selective university, students with three or more academic debts acquired during the first semester of study have a higher risk of dropping out compared to a moderately selective university. However, these differences were not significant in the regression models. In Russian universities, no differences were found depending on gender and field of study. In US universities, boys with the lowest level of academic performance (from 0 to 2 points) have a higher risk of dropping out compared to girls. The same trends are observed in engineering and science majors—students with the lowest performance levels have a higher risk of dropping out compared to students in other majors.

An increase in GPA between the first and second study periods reduces the risk of dropping out in all models considered. The heterogeneity of the influence of this indicator is manifested only in the strength of the connection: the effect is somewhat lower in the sample of boys and the sample of technical and natural science majors in US universities, as well as in the sample of girls and in the sample of selective university students in Russian universities.

C. The academic workload of students at American universities (number of credits).

With regard to this indicator, there are clear differences between the groups of students

under consideration. Firstly, according to the results of regression analysis, the volume of the academic load does not affect the dropout of male students, representatives of technical and natural science specialties, as well as those students who have not determined their field of study by the first academic period. Secondly, heterogeneity manifests itself in relation to universities of various types of selectivity. Thus, in universities with a low level of selectivity, only the highest load (17 credits or more) significantly reduces the risk of dropping out, and the average load (15-17 credits) does not have a significant relationship. In universities of the average type of selectivity, both medium (15-17) and high levels of academic workload (17 or more credits) significantly reduce the risk of dropping out, compared with a low level of workload. The observed heterogeneity may be explained by institutional differences in academic year systems (semesters or trimesters), as well as the special requirements of specific study programs regarding the volume of teaching load in the study period under consideration.

D. Inclusion of adaptation courses in the curriculum in the first academic period.

The results of the descriptive analysis indicate that taking adaptation courses in the first academic period significantly increases the risk of dropping out, however, the strength of this relationship varies depending on the selectivity of the university, gender, and field of study. The results of assessing regression models confirm these trends for all subsamples studied, except for the subsample of students studying in non-technical and natural science majors (social sciences, humanities, etc.)

E. Declaration of field of study in the first academic period.

The results of regression analysis indicate that in highly selective universities, the risk of expulsion of students who did not choose a field of study for the first academic period does not differ significantly from the risk of expulsion of other students. However, in universities with a lower level of selectivity this indicator is significant - if a student has not chosen his field of study, he/she has a higher risk of expulsion.

Appendix 2

Table 1. Event-history analysis estimates of academic momentum measures on student attrition from the first institution of enrollment in Ohio and Russian universities during 2.5 years observation period

Ohio universities			Russian universities		
Variable	Odds Ratio	St. Error	Variable	Odds Ratio	St. Error
Delayed enrollment in college[1]	1.14	0.11	Delayed enrollment in college[1]	1.46***	0.24
Took at least one developmental course[2]	1.15***	0.05	Failed one course in the first term[3]	1.87***	0.17
Attempted credits	0.97***	0.01	Failed two courses in the first term[3]	3.05***	0.33
First-term cumulative GPA	0.41***	0.01	Failed three or more courses in the first term[3]	7.64***	0.76
Grade improved over time[4]	0.47***	0.02	Grade improved over time[4]	0.35***	0.03
No major reported[5]	1.01	0.07			
Constant	0.08***	0.04	Constant	0.002***	0.002
Other control variables	Yes		Other control variables	Yes	
Dummy for bi-monthly periods	Yes		Dummy for bi-monthly periods	Yes	
University fixed effects	Yes		University fixed effects	Yes	
Regional-level variables	Yes				
N	246036		N	46951	
Pseudo R-square	0.219		Pseudo R-square	0.172	

Notes: * p<.05, ** p<.01, *** p<.001. The table reports only coefficients on the measures of academic momentum. Table 1A with full coefficients can be found in the Appendix.
Control variables in the Ohio model: student's age at the time of enrollment, gender, race/ethnicity, student's housing arrangements during the first term, major selected at the time of enrollment, receipt of a need-based grant, a merit-based grant, a loan and work-study appointment during any given term.
Control variables in the Russian model: student's age at the time of enrollment, gender, residency status combined with university housing (residents of the city of university, city non-residents living in university dormitories and city non-residents living in non-university housing) during the first term, selected academic discipline at the time of enrollment, receipt of tuition subsidies from the government, and award of the gold medal for academic performance in high school.
The number of observations reflects the number of person-bi-monthly observations in the case of Ohio and the number of person-quarter observations in the case of Russia's institutions. Regional-level variables include the percentage of population with bachelor's degree or higher in student's area of residence and unemployment level in student's area of residence.
[1] Reference group: No delay in college enrollment
[2] Reference group: Did not take any developmental courses
[3] Reference group: Did not fail any courses in the first term
[2] Reference group: Grades worsened between first and second terms
[4] Reference group: Reported Social and Behavioral sciences as a major