

Проблема инфляции оценок: измерение и анализ

Андрей Марков, Наталия Шагас, Елена Мерекина,
Александра Москалева

Статья поступила
в редакцию
в феврале 2024 г.

Марков Андрей Романович — кандидат экономических наук, доцент кафедры мировой экономики экономического факультета, советник декана, сопредседатель учебно-методической комиссии экономического факультета, Московский государственный университет им. М.В. Ломоносова. E-mail: amarkov1953@gmail.com. ORCID: <https://orcid.org/0000-0002-2291-5573>

Шагас Наталия Леонидовна — кандидат экономических наук, доцент кафедры микро- и макроэкономического анализа экономического факультета, член учебно-методической комиссии экономического факультета, Московский государственный университет им. М.В. Ломоносова. E-mail: nshagas@gmail.com

Мерекина Елена Владимировна — кандидат экономических наук, доцент кафедры финансов и кредита экономического факультета, член управляющего совета магистерской программы «Финансовые рынки и институты», Московский государственный университет им. М.В. Ломоносова. E-mail: elena.bulakhova@yandex.ru. ORCID: <https://orcid.org/0000-0001-9866-0856>

Москалева Александра Андреевна — ассистент кафедры микро- и макроэкономического анализа экономического факультета, Московский государственный университет им. М.В. Ломоносова. Адрес: 119991 Москва, ГСП-1, Ленинские горы, 3-й новый учебный корпус, экономический факультет. E-mail: sascha-mosk@gmail.com. ORCID: <https://orcid.org/0000-0001-5191-8725> (контактное лицо для переписки)

Аннотация

Проведен анализ количественных и качественных характеристик системы оценивания знаний студентов в бакалавриате отечественного вуза для определения наличия или отсутствия инфляции оценок. Под инфляцией оценок подразумевается выставление учащимся более высоких отметок независимо от того, имело ли место повышение уровня академической успеваемости. Рассматриваются три типа инфляции оценок: статический, динамический и дифференцированный. Предложена авторская трактовка каждого типа и способы их оценки.

Эмпирическую основу исследования составляет собранный авторами массив данных об оценках по обязательным и элективным дисциплинам, выставленным студентам нескольких волн приема в вуз. В количественном анализе применяются статистические и эконометрические методы. Предложен инструмент ранжирования дисциплин по характеру смещения распределения оценок, т.е. по степени завышения (инфляции) оценок, — аналог коэффициента асимметрии Пирсона. Выявлено завышение оценок по элективным дисциплинам (статическая инфляция). Динамической и дифференцированной инфляции оценок не обнаружено. На основании материалов обсуждения результатов исследования с преподавателями рассматриваются варианты политики, направленной на снижение инфляции оценок: нормативное регулирование распреде-

ления оценок, выравнивание структуры оценок на дисциплинах разных типов и введение ограничений на выбор дисциплин студентами.

Статья может быть полезна для руководителей и администраторов высших учебных заведений и экспертов, анализирующих системные вопросы развития университетского образования, в частности организацию оценивания.

Ключевые слова инфляция оценок, смещение распределения оценок, оценивание знаний студентов, коэффициент асимметрии Пирсона, факторный анализ

Для цитирования Марков А.Р., Шагас Н.Л., Мерекина Е.В., Москалева А.А. (2024) Проблема инфляции оценок: измерение и анализ. *Вопросы образования / Educational Studies Moscow*, № 3 (2), сс. 211–242. <https://doi.org/10.17323/vo-2024-19712>

The Problem of Grade Inflation: Measurement and Analysis

Andrei Markov, Natalia Shagas, Elena Merekina,
Alexandra Moskaleva

Andrei R. Markov — Candidate of Sciences (PhD) in Economics, Associate Professor at the Department of World Economy, Economic Faculty; Advisor to the Dean and Co-Chair of the Academic Quality Commission, Moscow State Lomonosov University. E-mail: amarkov1953@gmail.com. ORCID: <https://orcid.org/0000-0002-2291-5573>

Natalia L. Shagas — Candidate of Sciences (PhD) in Economics, Associate Professor at the Department of Micro and Macro Analysis, Economic Faculty; Member of the Academic Quality Commission, Moscow State Lomonosov University. E-mail: nshagas@gmail.com

Elena V. Merekina — Candidate of Sciences (PhD) in Economics, Associate Professor at the Department of Finance and Credit, Economic Faculty; Member of “Financial Markets and Institutions” Masters’ Program Managing Council, Moscow State Lomonosov University. E-mail: elena.bulakhova@yandex.ru. ORCID: <https://orcid.org/0000-0001-9866-0856>

Alexandra A. Moskaleva — Assistant Professor at the Department of Micro and Macro Analysis; Moscow State Lomonosov University Economic Faculty. E-mail: saschamask@gmail.com. ORCID: <https://orcid.org/0000-0001-5191-8725>

Abstract The article includes qualitative and quantitative analysis of the undergraduate student assessment system features in one of the universities in Russia in order to define presence of grade inflation. Grade inflation is understood as assignment of higher grades irrespective of improvement in student academic performance. Three types of grade inflation are considered: static, dynamic and differential. The authors offer interpretation of these types and suggest their methods for measurement.

The authors use specially collected grade data base of grades for compulsory and elective courses received by several cohorts of students admitted to a public research University. Statistical and econometric methods are used for quantitative analysis. The authors propose course ranking based on the degree of grade distribution shift that can be interpreted as a relative measure of grade inflation that is similar to Pearson asymmetry coefficient. Distribution of grades for elective courses show signs of static grade inflation. Dynamic and differential grade inflation were not detected in the sample analyzed. The paper explores grade inflation control pol-

icy options that were part of the consultations with various groups of academic staff. These include: normative regulation of grade distribution; equalizing of grade distribution for different types of courses; stricter regulation of student course selection by types of courses.

The paper could be useful for academic managers and higher education administrators and experts that study systemic issues in university education and assessment systems performance in particular.

Keywords grade inflation, grades distribution shifting; student assessment; Pearson asymmetry coefficient; grade inflation factor analysis

For citing Markov A.R., Shagas N.L., Merekina E.V., Moskaleva A.A. (2024) The Problem of Grade Inflation: Measurement and Analysis. *Voprosy obrazovaniya / Educational Studies Moscow*, no 3 (2), pp. 211–242 (In Russian). <https://doi.org/10.17323/vo-2024-19712>

1. Инфляция оценок как элемент внутривузовской экономики

В последние годы исследователи все чаще отмечают тенденцию к либерализации выставляемых студентам оценок: наблюдается поступательный рост доли отличных и хороших отметок как в целом, так и в разрезе разных групп студентов, групп дисциплин, направлений подготовки. В литературе и в обсуждениях все чаще звучит термин «инфляция оценок»¹ — настолько часто, что эксперты уже говорят о нем как о чем-то обыденном: «Привычное явление образования, известное как «инфляция оценок», происходит, когда оценки по курсу, субъективно выставляемые преподавателями, не соответствуют объективным показателям успеваемости учащихся» [Eiszler, 2002].

Инфляцию оценок достаточно давно обсуждают преподавательское сообщество и исследователи во многих странах. Ее отмечают как систематическое явление в колледжах и университетах США с 1960-х годов и начали активно исследовать с середины 1970-х [Juola, 1976]. Некоторые авторы связывают ее появление со стремлением студентов избежать призыва в армию во время войны во Вьетнаме и готовностью преподавателей им в этом помочь [Johnson, 2006]. Дополнительным стимулом для преподавателей выставять более высокие оценки стало распространение практики оценивания преподавателей студентами: эмпирические данные показывают, что студенты лучше оценивают те курсы, за которые они получают высокие оценки [Stroebe, 2020], и наблюдается устойчивая положительная связь между оценками, которые студенты выставляют преподавателям в ходе анкетирования после завершения учебного курса, и оценками, полученными студентами по дисциплинам [Eiszler, 2002].

¹ Аржанцев И.В. (2021) Почему не все студенты должны быть отличниками: <https://vc.ru/hse/252334-pochemu-ne-vse-studenty-dolzhen-byt-otlichnikami> (дата обращения: 20.08.2014); Revise Sociology (2023) UK Degrees and Grade-Inflation: <https://revisesociology.com/2023/06/26> (дата обращения: 20.08.2014).

К системным причинам инфляции оценок исследователи [Chowdhury, 2018; Finefter-Rosenbluh, Levinson, 2015] относят устойчивые различия в уровнях оценивания в дисциплинах разного типа: в естественных науках, с одной стороны, и в гуманитарных — с другой. Главное различие этих дисциплин с точки зрения оценивания знаний студентов состоит в том, что в естественных науках существует методическая возможность установить более ясные критерии выставления оценок, что существенно снижает издержки для преподавателей, если возникает необходимость в дополнительном обосновании выставленных баллов. Наличие сложившейся устойчивой дифференциации в критериях оценивания по дисциплинам разного типа, достоверных различий в жесткости оценивания и соответственно в уровне оценок по этим группам дисциплин подтверждено многочисленными исследованиями [Johnson, 2006]. Эти различия обнаружены как при анализе средних оценок, так и при исследовании распределений оценок в разных университетах, в частности в Университете Дюка, Университете штата Мичиган.

Наиболее широко в литературе представлен анализ инфляции оценок в университетах США, ее проявления и причины обсуждаются и в работах отечественных авторов [Павлова, 2009; Пискунов, Филимонова, 2006]. В последние годы появляются исследования, основанные на данных по России. В частности, на данных 18 онлайн-курсов УрФУ, размещенных на Национальной платформе открытого образования, показано, что онлайн-курсы не позволяют оценить реальный уровень знаний и навыков студентов, в результате чего возникает неконтролируемая инфляция оценок в онлайн-обучении [Ларионова, Гончарова, Дайнеко, 2022]. Таким образом, первые шаги в исследовании проблемы инфляции оценок в России сделаны, но эту работу необходимо продолжать, в том числе с использованием количественных методов.

Отношение к инфляции оценок в преподавательском сообществе неоднозначное. Некоторые считают ее естественным процессом, обусловленным тем, что высшее образование перестало быть элитарным и приобрело массовый характер. Массовизация высшего образования вызвана как потребностями современной экономики, так и стремлением государства бороться с неравенством в доходах, создавая социальные лифты. Однако значительное число преподавателей и экспертов высказывают озабоченность тенденцией к инфляции оценок, росту доли высших оценок и снижению их дифференциации. Такая озабоченность вполне понятна, поскольку с инфляцией оценок связан целый ряд сложностей и проблем в вузах: ослабление у студентов мотивации к дополнительным усилиям в обучении в условиях систематического роста доли высоких оценок; изменение поведения студентов в результате искажения стимулов к выбору тех или иных дисциплин и

формирование таким образом индивидуальных образовательных траекторий, которые не в полной мере соответствуют способностям и реальным знаниям учащихся.

Так, на основании анализа данных о выборе курсов для изучения, который совершали студенты девяти американских колледжей и университетов в течение 25 лет, исследователи [Sabot, Wakeman-Linn, 1991] пришли к выводу о существовании противоречия между системой стимулов, которые воздействуют на поведение студентов, и институциональными задачами расширения масштабов образования в области естественных наук и математики. Авторы использовали данные об изменении средних оценок в этих колледжах и университетах за период с 1962/1963 по 1985/1986 учебный год и показали, что в силу инфляции оценок в последние 25 лет постоянно искажались стимулы к получению образования. Сложились группы департаментов, на которых выставляются более низкие средние оценки (химия, математика и экономика), и департаментов с более высокими средними баллами (философия, английский язык, психология, политология). В выборе студентами дисциплин для изучения проявились две основные тенденции: с одной стороны, зафиксировано смещение выбора в пользу дисциплин с высокими оценками, что противоречило попыткам руководства университетов увеличить число студентов, изучающих естественные науки и математику в рассматриваемый период, с другой — отмечено снижение чувствительности студентов к полученным баллам на дисциплинах с более высоким уровнем оценок из-за малого их диапазона (всего четыре возможных варианта — 2, 3, 4, 5), т.е. из-за компрессии оценок. Авторы подчеркивают, что концентрированное распределение оценок на высоком уровне ослабляет их значимость как сигнала, которым они должны быть в системе образования.

В Университете Дюка в 1998/1999 учебном году осуществлен проект, в рамках которого также показано, что уровень выставляемых оценок оказывает существенное воздействие на выбор студентами дисциплин для изучения [Johnson, 2006]: обнаружено значительное снижение числа обучающихся, записывающихся на курсы по естественным наукам и математике, что, в свою очередь, подталкивает преподавателей к тому, чтобы повышать оценки, а студентов стимулирует выбирать дисциплины, которые ведут преподаватели, выставляющие оценки выше других.

Инфляция оценок также нарушает работу сигнальных механизмов, которые важны для работодателей при приеме на работу выпускников вузов. Фирмы не имеют возможности отличить действительно сильного выпускника от выпускника с такими же высокими оценками, которые получены у менее требовательных преподавателей. Оценки в дипломе перестают быть ориентиром, они выступают ложным сигналом.

Повышение общего уровня оценок ведет к росту самооценки у студентов вне зависимости от реальных учебных достижений. Довольно слабые студенты считают, что у них высокий уровень знаний, отсюда возникают неоправданные ожидания относительно будущей престижной работы с высокой оплатой, которые приводят к разочарованиям при столкновении с реальностью.

Преподаватели, стремящиеся выставлять оценки в соответствии со знаниями и способностями студентов к их усвоению, оказываются в невыгодном положении: на основании среднего балла в их группах можно сделать вывод, что преподаватель не способен обеспечить освоение студентами дисциплины на высоком уровне. Они оказываются вынуждены увеличивать долю хороших и отличных оценок за курс, т.е. попадают в «ловушку инфляции оценок».

Инфляция оценок оказывается устойчивым явлением при определенных условиях. Исследователи, проанализировавшие причины различий в выставлении оценок по естественным и гуманитарным дисциплинам в Университете Мичигана [Achen, Courant, 2009], установили, что уровень оценок представляет собой элемент внутривузовской экономики: он оказывает влияние на численность студентов по тем или иным дисциплинам и на размеры отдельных департаментов. Департаменты хотят видеть на своих дисциплинах умеренное число студентов, и при этом желательны сильные. Деканы колледжей заинтересованы в том, чтобы иметь достаточное число кредитных часов при равновесии между предложением дисциплин и спросом на них. Преподаватели, в свою очередь, стремятся выставлять такие оценки, которые они могли бы защитить без больших дополнительных затрат в случае жалоб студентов. Причем все эти факторы действуют в системе университетского образования, которая на протяжении как минимум пятидесяти лет функционирует в условиях инфляции оценок и в которой сложились устойчивые нормы и ожидания относительно уровня выставляемых баллов. Авторы отмечают, что дифференциация в распределении оценок по разным дисциплинам отражает рациональную реакцию преподавателей и администраторов на стимулы и процедуры, сложившиеся внутри университетов.

Инфляция оценок мешает проведению непредвзятого оценивания преподавателями и курсов студентами. В ходе проведения фокус-групп со студентами и преподавателями бакалавриата из шести вузов России выяснилось, что и те и другие высказывают опасения относительно связи инфляции оценок с высокими оценками, выставляемыми студентами преподавателям [Абрамова, Филькина, 2023].

Воздействие инфляции оценок на систему образования в целом весьма разнопланово. Выпускники могут оказаться менее

подготовленными, чем были бы в ее отсутствие, т.е. есть основания утверждать, что инфляция оценок может снижать уровень человеческого капитала и, следовательно, эффективность функционирования системы высшего образования.

В случае широкого распространения инфляции оценок осложняется работа органов управления системой образования и регулятора. Министерства и ведомства, отвечающие за образовательный процесс, могут воспринимать увеличение доли высоких оценок как индикатор повышения уровня образования, хотя на самом деле этот показатель свидетельствует, скорее, об обратном.

Приведенные данные о влиянии инфляции оценок на эффективность образовательного процесса однозначно указывают на необходимость систематического изучения проблемы оценивания, анализа динамики и структуры выставляемых оценок и факторов, определяющих их изменения. Важной частью осмысления процесса инфляции оценок является объективное измерение степени смещения оценок в сторону как завышения (инфляция), так и занижения (дефляция), что дает возможность проанализировать указанное явление как во времени, так и по его структуре.

Прямое соотнесение выставляемых оценок с уровнем освоения учебной дисциплины и оценка на этой основе степени инфляции оценок — задача достаточно сложная, и вряд ли ее можно решить напрямую. В каждом отдельном случае на существование смещения оценок в ту или иную сторону (в основном, конечно, в сторону завышения) могут указывать косвенные показатели объективности выставляемых отметок.

В частности, в настоящем исследовании для определения уровня инфляции оценок используется сравнение уровней средних оценок и показателей распределения оценок. Сравниваются доли разных оценок, выставляемых на разных дисциплинах одной и той же группе студентов; за одни и те же дисциплины в группах студентов последовательных лет поступления; в разных группах одного и того же года поступления. Полученные сравнительные характеристики позволяют судить о наличии и глубине инфляции оценок.

2. Стратегия исследования

В данной работе мы опираемся на классификацию, в соответствии с которой выделяются три типа инфляции оценок: статическая, динамическая и дифференцированная [Turner, Gershenson, 2020].

Статическая инфляция оценок отражает смысл выставляемых оценок в определенный момент времени, т.е. она отвечает на вопрос: на какой уровень владения материалом указывает та или иная оценка, соответствует ли уровень знаний оценке? Таким образом, статическая инфляция — это завышение успеваемости

учащихся относительно уровня знаний в определенный момент времени. Под статической инфляцией в данном исследовании понимается завышение оценок студентам одной и той же группы как по отдельным учебным дисциплинам, так и от первого семестра к последующим.

Динамическая инфляция оценок показывает, насколько изменился с течением времени уровень оценок относительно неизменного уровня знаний и навыков, демонстрируемых в рамках программ и курсов сходного наполнения. Именно это явление, по мнению авторов, имеют в виду большинство людей, когда говорят об инфляции оценок. Таким образом, динамическая инфляция — это изменение взаимосвязи между оценками и достижениями с течением времени. В данном исследовании динамическая инфляция — это повышение оценок по одним и тем же предметам у студентов последовательных лет набора.

Дифференцированная инфляция оценок имеет место, если взаимосвязь между оценками и достижениями различается для разных типов вузов или разных типов учащихся в статическом или динамическом разрезе или эта взаимосвязь меняется с разной скоростью. В понятии дифференцированной инфляции просматривается аналогия с концепцией инфляционного неравенства, согласно которой люди с разным уровнем дохода платят за основные товары разную цену: бедные — более высокую, богатые — более низкую. То есть дифференцированная инфляция — это вариативность взаимосвязи между оценками и достижениями в зависимости от типа вуза или у разных учащихся. В данной работе дифференцированная инфляция рассматривается как наличие различий в оценивании в зависимости от номера учебной группы на одном и том же курсе в случае, когда группы различаются по уровню первоначальной подготовки, т.е. этот уровень является одним из критериев формирования групп студентов на курсе.

Проведенный анализ научной литературы позволил выделить следующие основные направления количественного и качественного анализа инфляции оценок: документальное подтверждение самого тренда и его воздействия на выбор профиля обучения и учебных курсов [Sabot, Wakeman-Linn, 1991]; рассмотрение внутриуниверситетских институциональных механизмов воспроизводства инфляции, связанных с функционированием университета как организации [Achen, Courant, 2009]; более глубокий анализ влияния уровней оценивания на изменение поведения студентов, преподавателей и вузов на основе статистического и эконометрического аппарата [Chan, Hao, Suen, 2007].

Обнаружить в литературе количественный критерий для сравнительного анализа степени смещенности распределения оценок в разных учебных дисциплинах не удалось. Такой критерий нужен для того, чтобы сформировать объективную основу для вы-

явления инфляции оценок, измерения ее уровня и анализа факторов ее формирования. Без такого анализа не представляется возможной разработка системы конкретных мер по преодолению инфляции оценок. В настоящем исследовании делается попытка количественно оценить уровень инфляции оценок и тем самым заложить основу ее систематического анализа.

Мы последовательно оцениваем каждый из типов потенциальной инфляции оценок с использованием статистических и эконометрических методов на достаточно большой выборке данных. Сначала определяем наличие статической инфляции при помощи впервые предлагаемого показателя — аналога коэффициента асимметрии Пирсона, а также с использованием факторного анализа исследуем наличие связи между уровнем оценок на элективных дисциплинах и выбором студентами этих дисциплин. Далее выявляем характеристики дисциплин, уровень оценивания по которым наиболее высок.

На втором этапе мы выясняем, имеет ли место динамическая инфляция оценок: есть ли повышательный тренд в уровне оценок у студентов разных лет набора.

На третьем этапе с использованием регрессионного анализа мы тестируем гипотезу о наличии дифференцированной инфляции, т.е. различий в оценивании студентов с разным начальным уровнем знаний.

После проведения количественного анализа мы обсуждаем полученные результаты и связанные с ними мероприятия по снижению инфляции оценок, которые могут выбирать университеты.

3. Данные Ввиду отсутствия единого источника информации обо всех оценках и активностях студентов потребовалось разработать подходы к формированию базы данных оценок, пригодной для дальнейшей обработки. На подготовительном этапе исследования собраны данные нескольких волн приема в крупный столичный классический университет по направлениям подготовки «38.03.01 Экономика» и «38.03.02 Менеджмент» — с 2010 по 2019 г. В базу вошли 50 921 итоговая оценка по 127 дисциплинам — по всем предметам, по которым выставляется оценка, за исключением дисциплин с оценкой «зачет/незачет», практик и факультативных дисциплин. Итоговая выборка содержит оценки 2699 студентов, в том числе отчисленных и ушедших в академический отпуск, а также вышедших из академического отпуска на исследуемый курс обучения.

Для сбора данных использованы списки поступивших на первый курс, списки выпускников, списки распределения студентов по академическим группам, учебные карточки студентов со всеми итоговыми оценками по дисциплинам, данные о нагруз-

ке преподавателей, сканы ведомостей с оценками по дисциплинам, актуальные на год приема учебные планы, рабочие программы и календарно-тематические планы по дисциплинам, данные по студентам на момент поступления в вуз, результаты самообследования кафедр.

На первом этапе исследования сформированы выборки данных: по типам дисциплин (обязательные/вариативные), по кафедрам, по академическим группам, по лекторам и семинаристам, по направлению обучения («экономика»/«менеджмент»). Для расчетов использовались оценки студентов, полученные на сессии, т.е. до пересдач при их наличии.

На втором этапе выдвинуты гипотезы относительно причин выявленной смещенности оценок. Содержательные характеристики дисциплин представлены в рабочих программах дисциплин и календарных планах. К ним относятся доли баллов за групповую работу, за контрольные работы, за экзамен в итоговой сумме баллов на дисциплине, а также тип экзамена (устный/письменный) и предполагает ли дисциплина использование количественных методов. Регрессионные модели построены в программе *Gretl*.

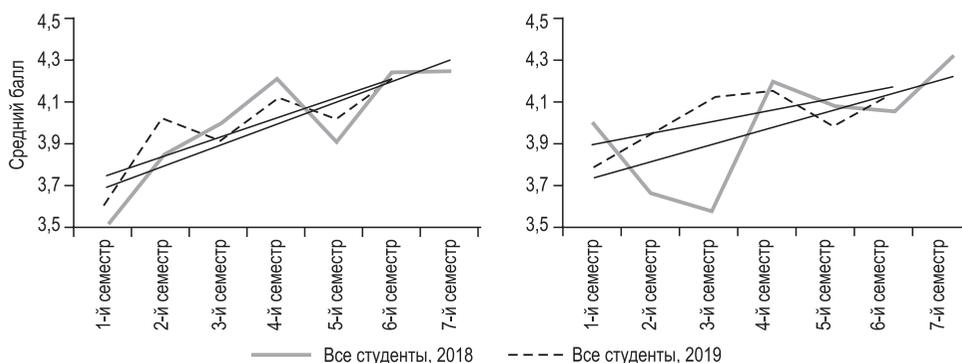
4. Методология определения наличия инфляции и ее применение
4.1. Статическая инфляция оценок: смещенность распределения оценок
4.1.1. Методика оценки смещенности распределения оценок

Для анализа статической инфляции сформирована подвыборка оценок, полученных студентами направлений «экономика» и «менеджмент» наборов 2018 и 2019 гг. за весь период завершеного обучения. Здесь мы анализируем наличие смещенности в распределениях оценок по отдельным дисциплинам или группам дисциплин. Группировка дисциплин осуществлялась, во-первых, по семестру обучения, во-вторых, по типу дисциплины: обязательная или вариативная.

Для предварительного анализа мы использовали динамику среднего балла по дисциплинам в разрезе семестров обучения (рис. 1).

На обоих направлениях наблюдается повышательный тренд в среднем балле, но на основании одного тренда мы не можем сделать вывод о том, присутствует ли смещенность в выставлении оценок. Чтобы определить наличие смещенности, нужно подобрать индикатор, который отражал бы завышение оценок на отдельной дисциплине или в группе дисциплин. В литературе в качестве такого индикатора используют, как правило, другую метрику оценки знаний, например, для школьников — итоговое тестирование [Gershenson, 2018]. В нашем случае показатель-аналог отсутствует, так как знания по всем предметам на выпуске не проверяются. Другой вариант индикатора смещенности — соответствие распределения нормальному закону. В случае фактически 4-балльной шкалы (оценки от 2 до 5 баллов) такой вариант тоже не представляется наилучшим, так как является слишком упрощен-

Рис. 1. Средний балл по всем предметам по семестрам обучения для студентов направлений «экономика» (слева) и «менеджмент» (справа)



Источник: Составлено авторами.

ным. Поэтому мы разработали собственную метрику определения смещенности — аналог коэффициента асимметрии Пирсона, который используется в статистике для определения скошенности распределения в сторону бóльших или меньших значений и может быть рассчитан по формуле:

$$As = \frac{\bar{x} - Mo}{\sigma}, \quad (1)$$

где \bar{x} — среднее значение; Mo — мода; σ — среднеквадратическое отклонение. Если значение коэффициента больше нуля, то распределение скошено в сторону более высоких значений, если меньше нуля — в сторону более низких.

Напрямую применить коэффициент асимметрии Пирсона не представляется возможным, так как зачастую мода распределения оценок по 4-балльной шкале составляет 4 балла и, соответственно, об асимметрии придется судить на основании превышения средним баллом студентов уровня 4 баллов. Поэтому мы предлагаем использовать аналог коэффициента асимметрии Пирсона, где вместо моды фигурирует некоторый «образцовый» показатель — бенчмарк, превышение которого свидетельствует о завышении уровня оценок на дисциплине. Таким образом получаем формулу расчета:

$$K = \frac{\bar{x} - benchmark}{\sigma}. \quad (2)$$

4.1.2. Выбор бенчмарка для оценки смещенности распределения оценок

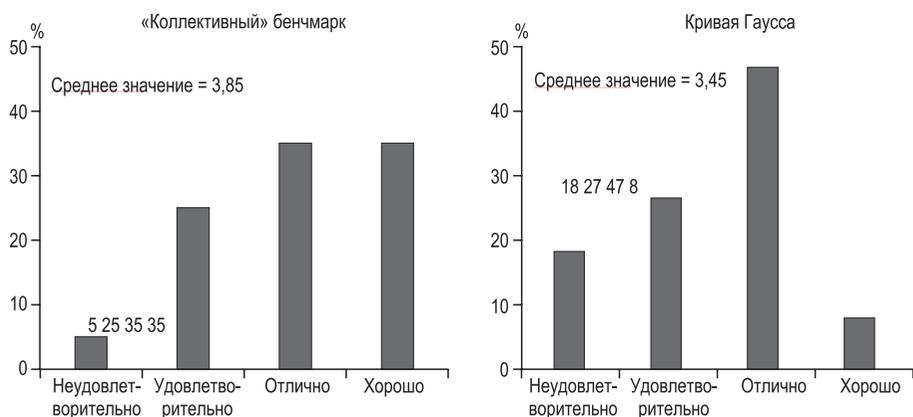
Ключевым вопросом в применении этого аналога коэффициента Пирсона является выбор бенчмарка или набора бенчмарков. В данном исследовании мы предлагаем использовать в качестве критерия смещенности распределения среднее значение за весь период обучения по результатам обучения на обязательных дисциплинах, так называемый коллективный бенчмарк. К достоин-

ствам показателя относятся, во-первых, то, что он учитывает уровень оценок всех студентов потока, во-вторых, что он включает оценки по курсам, прослушанным студентами на разных годах обучения. В нашей выборке, созданной для статического анализа, у студентов набора 2018 г., завершивших обучение на момент проведения анализа, этот показатель составил 3,85 для направления «экономика» и 3,9 — для направления «менеджмент».

Рассчитанный с использованием коллективного бенчмарка аналог коэффициента асимметрии Пирсона служит инструментом ранжирования дисциплин, который отличается от ранжирования по среднему значению тем, что учитывает разброс в распределении оценок благодаря делению на стандартное отклонение. Интерпретируется коэффициент следующим образом: если его значение больше нуля, то распределение оценок считается смещенным по отношению к коллективному бенчмарку в сторону более высоких оценок, если значение меньше нуля — напротив, в меньшую сторону.

В то же время называть распределение с положительным коэффициентом распределением с завышенным уровнем оценок было бы слишком поспешно, так как деление на основе коллективного бенчмарка является бинарным. Требуется определить критерии для выявления «несмещенных» дисциплин. Для этого мы ввели границы для бенчмарка, внутри которых распределения относятся к несмещенным. В качестве верхней границы выбрана мода распределения оценок — 4, в качестве нижней границы — среднее значение для оценок, распределенных по нормальному закону, которое составляет 3,45. Это значение подобрано по аналогии с распределением по кривой Гаусса, которое используется для распределений оценок при статистическом анализе

Рис. 2. Примеры распределений, соответствующих коллективному бенчмарку (слева) и нормальному закону (справа)



Источник: Составлено авторами.

качества преподавания и выставления оценок в Лиге академической честности Республики Казахстан², при помощи перевода из 5-балльной шкалы в 4-балльную. На рис. 2 приведены примеры распределений для коллективного бенчмарка и нормального распределения.

Предложенный коэффициент является прокси-измерителем инфляции оценок, так как напрямую измерить отклонение выставленных оценок от соответствующего уровня знаний на доступных данных не представляется возможным.

4.1.3. Пример использования предложенного показателя смещенности распределения оценок

На основе использования предложенных порогов все дисциплины, которые изучали студенты, набранные в вуз в 2018 г., были разделены на три категории: дисциплины со смещением оценок вверх (потенциально инфляционные) — для них аналог коэффициента Пирсона положителен при всех вариантах бенчмарка; несмещенные дисциплины и дисциплины со смещением вниз распределением оценок. Доли дисциплин каждого типа по направлениям приведены в табл. 1.

Таблица 1. Доли дисциплин с разным типом смещенности распределения оценок по направлениям подготовки (%)

	Экономика	Менеджмент
Смещены вверх	60	58
Не смещены	31	36
Смещены вниз	9	7

Источник: Составлено авторами.

Существенных различий между направлениями подготовки по долям дисциплин с разным типом смещенности не наблюдается, преобладают дисциплины с завышением среднего уровня оценок выше моды — 4 балла. Чтобы определить, какие факторы приводят к завышению оценок, мы выяснили, какие типы дисциплин попадают в группу с завышенным баллом. Для обязательных дисциплин аналог коэффициента Пирсона практически равен нулю по построению. Для элективных дисциплин рассчитан аналог коэффициента Пирсона с коллективным бенчмарком. Для дисциплин вариативной части учебного плана направления «экономика» получено значение 0,29, для направления «менеджмент» — 0,45. Таким образом, наблюдается относительное завышение уровня оценок на дисциплинах по выбору по сравнению с обязательными

² Критерии оценки результатов обучения студентов на основе общих принципов Лиги академической честности: <https://adaldyq.kz/> (дата обращения: 20.08.2024).

ми дисциплинами, при этом на направлении «менеджмент» это завышение более сильное.

4.2. Статическая инфляция оценок: выбор дисциплин студентами

Далее была предпринята попытка оценить наличие у студентов «стратегического» поведения, при котором обучающиеся снижают себе нагрузку, опираясь при выборе дисциплин на отзывы студентов прошлых лет обучения о том, насколько сложно сдать тот или иной предмет. Деление предметов на обязательные и вариативные, формирование треков обучения и другие схожие практики, с одной стороны, позволяют учесть научные и карьерные интересы учащихся, но, с другой стороны, могут обусловить появление конкуренции за студентов между дисциплинами по выбору. Она возникает, если, например, студент обязан выбрать несколько предметов из списка или же ставится условие о запуске только тех дисциплин, на которые записалось определенное число слушателей. Преподаватели вариативных предметов могут быть заинтересованы в более лояльном оценивании, чтобы их дисциплины выбирали в будущем, так как от этого зависит уровень их нагрузки и оплаты труда.

Выбирая дисциплины для изучения, студенты предпочитают наиболее интересные для них предметы и те, заниматься которыми они мотивированы в силу учебного плана и в расчете на академическую перспективу — например, выбирают дисциплины с той же кафедры, где пишут исследовательские работы. Однако при выборе дисциплин или треков обучения они могут обращать внимание и на ожидаемый уровень оценок, о котором узнают от студентов старших курсов или из результатов анкетирования, если они публикуются. Для выявления этого эффекта построена регрессионная модель по данным о дисциплинах по выбору для студентов 2016–2019 гг. набора. В общем виде модель описывается уравнением:

$$pct_student_choose_{it} = \alpha_1 + \alpha_2 \cdot grade_level_{it-1} + \alpha_3 \cdot X_{it} + \varepsilon_{it}, \quad (3)$$

где $pct_student_choose_{it}$ — доля студентов, записавшихся на дисциплину; $grade_level_{it}$ — переменные интереса, которые описывают уровень оценок на дисциплине с лагом в один год: коэффициент Пирсона, средний балл студентов на дисциплине и доля оценок 4 и 5 в общем числе оценок; X_{it} — матрица контрольных переменных (доля баллов за групповую работу в сумме баллов за дисциплину; оценка студентами дисциплины в предыдущем году по итогам анкетирования; доля студентов, которые выбрали дисциплину и пишут работу на соответствующей кафедре, среди всех студентов, которые пишут курсовую на этой кафедре; логарифм числа альтернатив в семестре — для учета семестра и на-

правления обучения); α_1 — константа; α_2, α_3 — оценки коэффициентов; ε_{it} — случайная ошибка. Результаты оценки представлены в табл. 2.

Таблица 2. Результаты оценки регрессий для выявления наличия связи между уровнем оценивания и выбором студентами дисциплины

Модели	(1)	(2)	(3)
Константа	-0,13 (0,39)	0,34 (0,25)	0,08 (0,27)
Ln (число альтернативных дисциплин для выбора)	-0,07 (0,05)	-0,06 (0,05)	-0,073* (0,043)
Доля баллов за групповую работу	0,54** (0,23)	0,58** (0,23)	0,54** (0,22)
Доля записавшихся на дисциплину курсовиков кафедры	0,22** (0,1)	0,21* (0,11)	0,22** (0,1)
Оценка дисциплины по итогам анкетирования, $t - 1$	0,008 (0,06)	0,014 (0,06)	0,008 (0,05)
Средняя оценка студентов, $t - 1$	0,13 (0,08)		
Аналог коэффициента Пирсона, $t - 1$		0,04 (0,06)	
Доля оценок 4 и 5 в общем числе оценок, $t - 1$			0,42** (0,19)
Число наблюдений	115	115	115
R^2_{adj}	0,11	0,09	0,13

Примечание: Зависимая переменная – $pct_student_choose_{it}$. В скобках указаны робастные стандартные ошибки; * значимость на 10%-ном уровне; ** на 5%-ном уровне; *** на 1%-ном уровне.

Источник: Составлено авторами.

Ввиду необходимости использовать лаги для оценки связи уровня отметок с выбором студентами отдельных дисциплин, включенности в анализ периода дистанционного обучения, а также наличия сильной инертности в выборе студентами дисциплин с опорой на опыт предшествующих поколений размер выборки и объясняющая сила моделей недостаточны для однозначного вывода о существовании рассматриваемой зависимости. Однако представленный вариант моделирования иллюстрирует подход к анализу стратегического поведения студентов при выборе дисциплин. Также получен логичный результат: относительные метрики, не наблюдаемые студентами в явном виде, такие как смещенность относительно бенчмарка и средний уровень оценок, по всей видимости, не так существенны для выбора. В то же время получен значимый на 5-процентном уровне положительный ко-

эффицент при доле оценок 4 и 5 на дисциплине. Что логично, так как такая метрика распределения оценок легко оценивается «на глаз» и может быть передана в коммуникации между студентами разных курсов.

4.3. Статическая инфляция оценок: факторы завышения уровня оценок

На следующем этапе статического анализа мы выдвинули предположение, что различия в уровнях оценивания связаны с характеристиками дисциплин, и перешли от анализа общих тенденций, наблюдаемых в данных, к поиску факторов получения более высоких оценок на микроуровне — на уровне характеристик отдельных дисциплин.

Для выявления факторов завышения оценок по отдельным дисциплинам использована пространственная выборка, состоящая из оценок студентов, характеристик студентов и характеристик дисциплин для обоих направлений подготовки за 2018 и 2019 гг. набора. Чтобы выделить значимые характеристики дисциплин, которые здесь выступают в роли переменных интереса, мы оценили множество регрессионных уравнений методом МНК с робастными стандартными ошибками и контролем на характеристики студента, год набора, направление подготовки и кафедральные особенности. В общем виде модель может быть записана следующим образом:

$$Grade_{it} = \alpha_1 + \alpha_2 \cdot Z_{it} + \alpha_3 \cdot X_{it} + \mu_t + \varepsilon_{it}, \quad (4)$$

где $Grade_{it}$ — оценка, полученная студентом на сессии; Z_{it} — матрица переменных интереса (характеристики дисциплин); X_{it} — матрица контрольных переменных (характеристики студента (олимпиады, регионы, результат ЕГЭ по математике, фиктивные переменные групп, фиктивные переменные кафедр); α_1 — константа; α_2, α_3 — векторы оценок коэффициентов; μ_t — фиктивные переменные года/семестра обучения; ε_{it} — случайная ошибка.

Описательные статистики и детализация использованных характеристик дисциплин и студентов приведены в табл. 3.

Тестировались следующие гипотезы относительно связи характеристик дисциплин с оценками.

1. На гранулярных данных будет подкреплён вывод предварительного анализа о том, что оценки по вариативным дисциплинам в среднем выше, чем по обязательным. Эта гипотеза подтвердится, если при фиктивной переменной вариативного типа дисциплины будет получен положительный коэффициент.
2. Высокая доля групповой работы будет оказывать повышающее давление на уровень оценок ввиду наличия проблемы безбилетника и сложности дифференцированного оце-

Таблица 3. **Описательные статистики переменных модели для определения значимых характеристик дисциплин, на которых выставляются более высокие оценки**

	Название	Описание	Среднее	Минимум	Максимум
	Оценка	Оценка по дисциплине	4,04	2	5
Характеристики дисциплин	Дисциплина по выбору	= 1, если дисциплина по выбору студента	0,37	0	1
	Доля баллов за групповую работу	Доля ³ баллов за коллективные работы (кейсы, проекты)	0,06	0	0,56
	Доля баллов за индивидуальную аудиторную работу	Доля баллов за индивидуальные работы в аудитории (контрольные, тесты)	0,28	0	0,7
	Доля баллов за индивидуальную домашнюю работу	Доля баллов за индивидуальные работы дома (домашние задания, эссе, рефераты)	0,08	0	0,6
	Доля баллов за работу на семинарах	Доля баллов за посещение или работу на семинарах	0,22	0	0,85
	Доля баллов за экзамен	Доля баллов за экзамен	0,36	0	0,67
	Устный экзамен	= 1, если экзамен устный	0,41	0	1
	Инструментальная дисциплина	= 1, если дисциплина инструментальная ⁴	0,41	0	1
Характеристики студента	Балл за ЕГЭ по математике	Балл за ЕГЭ по математике	78,1	0	100
	Иностранец	= 1, если студент – иностранец	0,095	0	1
	Без вступительных испытаний	= 1, если студент поступил по льготе «без вступительных испытаний»	0,02	0	1
	Студент из региона	= 1, если студент не из Москвы	0,39	0	1
	Направление «экономика»	= 1, если студент учится на направлении «экономика»	0,66	0	1

Источник: Составлено авторами.

- нивания вклада отдельных членов команды. Эта гипотеза подтвердится, если коэффициент при доле групповой работы в общем количестве баллов за дисциплину окажется положительным.
3. Высокая доля контрольных работ в аудитории будет отрицательно сказываться на оценках по причине наличия возможностей наблюдения за выполнением задания. Подтверждение гипотезы предполагает отрицательную связь между оценками и долей контрольных работ в структуре оценки по дисциплины.
 4. Устный формат экзамена будет повышать уровень оценивания, так как, с одной стороны, у преподавателя есть возможность задать студенту дополнительные вопросы и тем самым улучшить качество ответа, а с другой — в условиях устного экзамена сильнее проявляется влияние личных отношений преподавателя и студента при отсутствии строгих критериев оценивания таких испытаний. О подтверждении

³ Здесь и далее доля в максимальном количестве баллов за дисциплину.

⁴ Предполагает использование и изучение математических методов, моделей, программирования.

этой гипотезы может свидетельствовать положительный коэффициент при фиктивной переменной, отвечающей за устный формат экзамена по дисциплине.

Для проверки устойчивости результатов построены три спецификации модели, при этом использованы три варианта учета контрольных переменных: первый — базовый, без включения дополнительного контроля, второй — с временными эффектами года обучения и семестра (рассматриваемый промежуток времени включает часть периода дистанционного обучения вследствие пандемии COVID-19), третий — с учетом кафедральных особенностей посредством включения $(n - 1)$ фиктивной переменной кафедр. Во все модели дополнительно в качестве контроля включены номера группы от 1 до 12 (чем меньше номер группы, тем выше в среднем баллы, полученные студентами на распределительных тестах в начале обучения).

Чтобы избежать проблемы мультиколлинеарности, в модели не включается переменная доли баллов за домашние работы студента. Кроме того, так как между результатом ЕГЭ по математике и номером группы на направлении «экономика» наблюдается корреляция $(-0,51)$, проведен тест на мультиколлинеарность при помощи VIF-коэффициентов. Коэффициенты не превысили пороговое значение 10 — следовательно, проблема мультиколлинеарности не влияет на оценку. Результаты оценки спецификаций представлены в табл. 4.

Таблица 4. Результаты оценки регрессий для выявления наличия характеристик дисциплин, значимых для получения более высоких оценок

Модели	(1)	(2)	(3)
Константа	3,9*** (0,07)	3,5*** (0,07)	3,4*** (0,1)
Дисциплина по выбору	0,04*** (0,01)	-0,01 (0,01)	0,07*** (0,02)
Направление «экономика»	0,27*** (0,02)	0,25*** (0,02)	0,24*** (0,02)
Доля баллов за групповую работу	0,39*** (0,06)	0,53*** (0,056)	0,53*** (0,065)
Доля баллов за индивидуальную аудиторную работу	-0,59*** (0,05)	-0,4*** (0,055)	-0,51*** (0,07)
Доля баллов за работу на семинарах	-0,06 (0,05)	0,12** (0,052)	-0,11* (0,062)
Доля баллов за экзамен	-0,32*** (0,08)	-0,19** (0,085)	0,25** (0,11)
Экзамен устный	0,15*** (0,01)	0,16*** (0,012)	0,1*** (0,017)

Окончание табл. 4

Модели	(1)	(2)	(3)
Инструментальная дисциплина	-0,36*** (0,02)	-0,32*** (0,016)	-0,25*** (0,025)
Балл за ЕГЭ по математике	0,007*** (0,0006)	0,008*** (0,0006)	0,007*** (0,0006)
Иностраный гражданин	0,06** (0,02)	0,07*** (0,02)	0,08*** (0,02)
Без вступительных испытаний	-0,02 (0,04)	0,009 (0,04)	0,001 (0,04)
Студент из региона	0,12*** (0,01)	0,12*** (0,01)	0,13*** (0,01)
Номер группы	+	+	+
Номер семестра обучения		+	+
Фиксированные эффекты года		+	+
Фиксированные эффекты кафедры			+
Число наблюдений	24 232	24 232	24 232
R^2_{adj}	0,15	0,16	0,19

Примечание: Зависимая переменная – *grade*. В скобках указаны робастные стандартные ошибки; * значимость на 10%-ном уровне; ** на 5%-ном уровне; *** на 1%-ном уровне.

Источник: Составлено авторами.

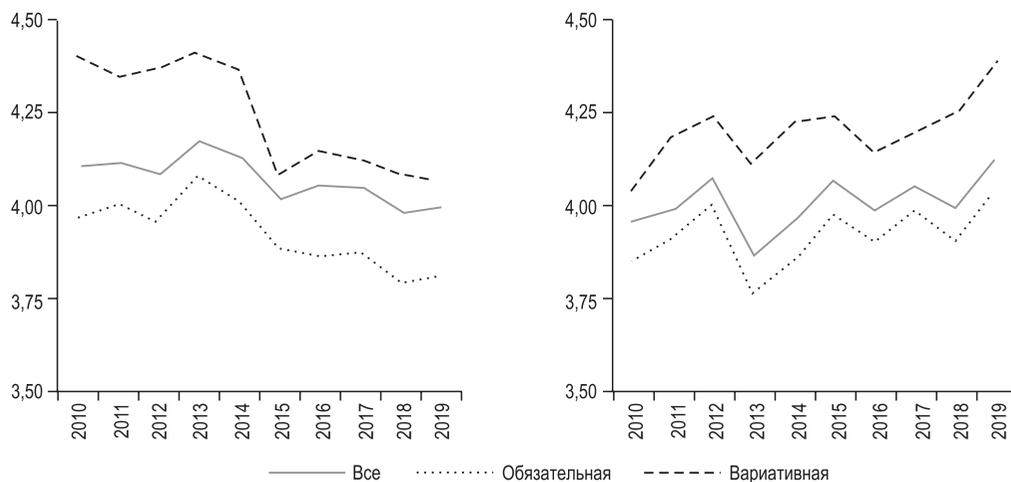
В результате регрессионного анализа получен устойчивый относительно предложенных вариантов контроля результат, свидетельствующий о положительном влиянии на уровень оценивания студентов более высокой доли групповой работы, более низкой доли контрольных работ и наличия устного экзамена, что согласуется с нашими предположениями. В то же время регрессии показывают неустойчивость знака при переменных доли баллов за работу на семинарах и баллов за экзамен при добавлении контроля на кафедру. Этот результат предположительно связан с кафедральной спецификой оценивания данных видов работ и требует дополнительного анализа.

4.4. Динамическая инфляция оценок

Для анализа инфляции оценок в динамике, т.е. для выявления наличия или отсутствия систематического завышения уровня оценивания студентов от набора к набору, мы использовали оценки подвыборки студентов из четырех групп в середине потока для обоих направлений обучения за 10 лет — с 2010 по 2019 год набора. Середина потока выбрана потому, что студенты распределены по группам по начальному уровню знаний — а значит, середина потока наиболее репрезентативна для иллюстрации динамики оценок набора. Создание таких подвыборок позволяет избежать больших трудозатрат на ручной сбор данных по всем студентам.

Для подвыборок рассчитаны средние значения по всем оценкам, полученным студентами этого года набора, по всем дисциплинам в целом и отдельно по типам дисциплин. Полученные результаты представлены на рис. 3.

Рис. 3. Динамика среднего балла по годам набора на направлениях «экономика» (слева) и «менеджмент» (справа)



Источник: Составлено авторами.

Явного повышательного тренда в оценивании студентов на рассматриваемом периоде не наблюдается. Наблюдаемые тренды характерны как для всех оценок, так и для подгрупп обязательных и вариативных дисциплин. В то же время превышение уровня оценивания на дисциплинах по выбору оказалось устойчивым трендом и наблюдается на всем 10-летнем промежутке. Таким образом, на этапе предварительного анализа предпосылок наличия динамической инфляции не выявлено, правда, только если считать неизменным качество, или уровень подготовленности, поступающих студентов по годам набора.

4.5. Дифференцированная инфляция оценок: базовая спецификация

Анализ дифференцированной инфляции предполагает выявление различий в оценивании групп студентов, различающихся по какому-либо общему признаку. В нашем исследовании мы решили оценить наличие дифференцированной инфляции на основании различий в оценивании в зависимости от уровня первоначальной подготовки студентов. Предположительно может наблюдаться как отсутствие этих различий, если ко всем студентам применяется единая четко прописанная балльно-рейтинговая система и единые критерии оценки разных видов работ, так и наличие различий,

если преподаватели оценивают студентов в группе относительно друг друга, а не по единым критериям для потока.

Для анализа дифференцированной инфляции оценок модель, использованная нами в статическом анализе, рассмотрена под другим углом: теперь в качестве переменной интереса выступают номер группы, в которой обучается студент, и балл ЕГЭ по математике. Гипотеза, которую мы закладываем в предлагаемую модель, следующая: если ни номер группы, ни результат ЕГЭ не являются значимыми характеристиками уровня оценок, то первоначальные способности и знания, которые аппроксимируют эти показатели, перестают быть существенны для результатов обучения. Такой результат оценки моделей может служить как показателем наличия дифференцированной инфляции, так и индикатором того, что уровень способностей и знаний студентов в процессе обучения сильно меняется по сравнению с первоначальным. Следовательно, может потребоваться разработка дополнительных подходов к определению причин отсутствия различий.

Для выявления дифференцированной инфляции к моделям, представленным в разделе статического анализа, добавлены переменные наклона для перемножения фиктивных переменных семестра на номер группы. Если перед номером группы сохраняется значимый отрицательный коэффициент для всех семестров обучения — значит, дифференцированной инфляции нет, студенты с разным уровнем первоначальной подготовки оцениваются по-разному.

Для проверки устойчивости результатов построены модели на сокращенном временном промежутке: начиная с 3-го семестра (2-й курс обучения). Эти модели позволяют избежать возможного сильного влияния более низких оценок по математическим предметам на 1-м курсе, обусловленных трудностями адаптации студентов к самостоятельному обучению в университете и повышенной сложностью этих предметов для многих выпускников школ. Результаты базовых и дополнительных спецификаций приведены в табл. 5.

Таблица 5. Результаты оценки регрессий для выявления наличия различий в оценках, получаемых студентами разных групп

	Все студенты, 1–7-й семестры	Все студенты, 1–7-й семестры	Все студенты, 3–7-й семестры
Константа	3,5*** (0,11)	3,7*** (0,11)	3,5*** (0,2)
Направление «экономика» = 1	0,23*** (0,02)	0,23*** (0,02)	0,23*** (0,02)
Балл ЕГЭ по математике	0,007*** (0,0006)	0,007*** (0,0006)	0,006*** (0,0007)

Окончание табл. 5

	Все студенты, 1–7-й семестры	Все студенты, 1–7-й семестры	Все студенты, 3–7-й семестры
№ группы	–0,06*** (0,002)	–0,096*** (0,005)	–0,05*** (0,006)
№ группы*(Семестр = 2)		0,03*** (0,007)	
№ группы*(Семестр = 3)		0,05*** (0,007)	
№ группы*(Семестр = 4)		0,037*** (0,006)	–0,009 (0,007)
№ группы*(Семестр = 5)		0,036*** (0,007)	–0,018*** (0,007)
№ группы*(Семестр = 6)		0,05*** (0,007)	0,004 (0,007)
№ группы*(Семестр = 7)		0,049*** (0,007)	0,0008 (0,008)
Характеристики дисциплин	+	+	+
Характеристики студентов	+	+	+
Фиксированные эффекты семестра	+	+	+
Фиксированные эффекты года обучения	+	+	+
Фиксированные эффекты кафедр	+	+	+
Число наблюдений	24 232	24 232	17 057
R^2_{adj}	0,19	0,19	0,14

Примечание: Зависимая переменная – *grade*. В скобках указаны робастные стандартные ошибки; * значимость на 10%-ном уровне; ** на 5%-ном уровне; *** на 1%-ном уровне.

Источник: Составлено авторами.

Номер группы и результат ЕГЭ оказываются значимыми для любых спецификаций моделей, что свидетельствует об отсутствии существенной конвергенции студентов по уровню оценок на старших курсах. Студенты с более низким уровнем первоначальных знаний продолжают в среднем учиться хуже, чем учащиеся, более подготовленные на этапе поступления. Таким образом, дифференцированная инфляция в наших данных отсутствует: оценивание связано с уровнем подготовки, о чем свидетельствует значимость его прокси-переменных. При этом для спецификации на всем периоде обучения наблюдается сокращение разрыва между студентами с разным первоначальным уровнем знаний: коэффициенты при переменных наклона положительны для всех семестров относительно 1-го семестра. На подвыборке оценок за 3–7-й семестры такой тенденции не наблюдается, поэтому можно

предположить, что основное сокращение разрыва в уровне первоначальных знаний происходит на 1-м курсе, в том числе посредством отсева наименее подготовленных студентов.

4.6. Дифференцированная инфляция оценок: подвыборки по направлениям подготовки

Далее рассмотрим наличие дифференцированной инфляции по направлениям подготовки. В отличие от «экономики» на «менеджменте» студенты не пишут распределительный тест по математике и разделяются на группы только на основе теста по английскому языку. С одной стороны, результат ЕГЭ по иностранному языку коррелирует с результатом ЕГЭ по математике и на направлении «менеджмент» меньше математических предметов, с другой стороны, отбор устроен по-разному. Результаты оценки представлены в табл. 6.

Таблица 6. Результаты оценки регрессий для выявления различий в оценках, получаемых студентами разных групп, по направлениям подготовки

	Экономика, 1–7-й семестры	Экономика, 3–7-й семестры	Менеджмент, 1–7-й семестры	Менеджмент, 3–7-й семестры
Константа	4,2*** (0,22)	4,4*** (0,28)	3,5*** (0,16)	2,6*** (0,19)
Балл ЕГЭ по математике	0,007*** (0,0009)	0,005*** (0,001)	0,01*** (0,0008)	0,009*** (0,0009)
№ группы	-0,094*** (0,006)	-0,049*** (0,007)	-0,097*** (0,01)	-0,14*** (0,012)
№ группы*(Семестр = 2)	0,04*** (0,008)		-0,02 (0,02)	
№ группы*(Семестр = 3)	0,043*** (0,008)		-0,036** (0,02)	
№ группы*(Семестр = 4)	0,035*** (0,007)	-0,007 (0,008)	-0,02 (0,02)	0,02 (0,02)
№ группы*(Семестр = 5)	0,033*** (0,008)	-0,02* (0,008)	-0,004 (0,02)	0,03 (0,02)
№ группы*(Семестр = 6)	0,06*** (0,007)	0,01* (0,008)	0,02 (0,02)	0,05*** (0,02)
№ группы*(Семестр = 7)	0,052*** (0,008)	0,005 (0,009)	0,071*** (0,02)	0,11*** (0,02)
Характеристики дисциплин	+	+	+	+
Характеристики студентов	+	+	+	+
Фиксированные эффекты семестра	+	+	+	+
Фиксированные эффекты года обучения	+	+	+	+
Фиксированные эффекты кафедр	+	+	+	+
Число наблюдений	15 801	11 329	8431	5728
R^2_{adj}	0,23	0,16	0,19	0,21

Примечание: Зависимая переменная – *grade*. В скобках указаны робастные стандартные ошибки; * значимость на 10%-ном уровне; ** на 5%-ном уровне; *** на 1%-ном уровне.

Источник: Составлено авторами.

Таким образом, результат совпадает для обоих направлений подготовки, несмотря на различия в способах распределения по группам. В то же время для направления «менеджмент» не характерно выравнивание уровня по группам по сравнению с первым семестром обучения. В нашей подвыборке выравнивание на направлении «менеджмент» наблюдается на 4-м курсе обучения.

**5. Дискуссия:
альтернативные
варианты
политики**

Администраторы системы высшего образования, экспертное и преподавательское сообщество уделяют значительное внимание проблеме инфляции оценок. Анализируются масштабы и истоки проблемы, формы ее проявления. Идет поиск управленческих и педагогических решений, направленных на то, чтобы контролировать процесс увеличения доли выдаваемых дипломов с отличием, рост средних баллов и повышение доли отличных оценок. В то же время вопрос о необходимости административного регулирования инфляции оценок остается дискуссионным. Далее мы рассмотрим возможные подходы к управлению оцениванием знаний.

**5.1. Системные
и институцио-
нальные подходы**

Проблему инфляции оценок можно рассматривать и решать с позиций двух основных подходов: системно-скоординированного и децентрализованного. В рамках системно-скоординированного подхода органы управления образованием системно собирают информацию и проводят диагностику состояния дел с инфляцией оценок в целом по системе образования или по большой группе вузов. На основе обобщенного анализа формулируются общие способы реагирования на выявленные тенденции. При децентрализованном подходе отдельные учебные заведения самостоятельно определяют методы диагностики и способы реагирования. Каждый из этих подходов имеет свои достоинства и недостатки.

Примером общесистемного подхода может служить работа Офиса студенческих дел [Office for Students, 2022], который является регулятивным и аккредитационным органом в системе высшего образования Великобритании. Реагируя на рост доли выдаваемых дипломов с отличием — по итогам 2021/2022 учебного года 32,8% выпускников-бакалавров получили отличные дипломы первого уровня, в то время как в 2010/2011 учебном году их было 15,5%, — Офис провел детальный сравнительный анализ факторов, объясняющих этот рост. Оказалось, что около половины прироста доли дипломов с отличием связано с характеристиками учебного заведения и изменением входного уровня знаний при поступлении на программу, но остальной прирост относится к «необъяснимым» факторам и рассматривается как показатель инфляции оценок. Помимо широкого обсуждения полученных результатов Офис инициировал расследование случаев наиболее

значительных отклонений выданных квалификаций от утвержденных стандартов. По сообщениям прессы, в случае подтверждения необоснованной выдачи дипломов с отличием вузам грозил штраф в размере до 0,5 млн фунтов стерлингов до 2% общих доходов университета⁵.

Один из самых известных примеров политики по снижению инфляции оценок на уровне отдельного университета — инициатива Принстонского университета 2004 г. Руководство университета приняло административное решение о том, что доля отличных оценок по отдельным дисциплинам не должна превышать 35%. В период действия этого правила средний балл снизился, но после его отмены в 2014 г. вернулся к прежним значениям и даже превзошел их. В Принстоне ожидали, что другие ведущие университеты США применят такую же или сходные меры, но этого не произошло. Политика вызывала недовольство преподавателей и особенно выпускников Принстона, которые жаловались, что они стали проигрывать на рынке труда и при отборе в магистратуры выпускникам других вузов, где подобные правила не применялись.

Вопрос об инфляции оценок и смещенности их распределения обсуждается в Гарвардском университете. Осенью 2023 г. ученому совету университета был представлен доклад, в котором фиксировалось значительное увеличение доли отличных оценок — с 60% в 2010/2011 году до 79% в 2020/2021-м, а также рост средних оценок по 4-балльной шкале за тот же период с 3,41 до 3,80. При разработке мер, которые могли бы снизить инфляцию оценок, рассматривался, в частности, вариант ослабить зависимость присуждения преподавательских премий от результатов студенческого анкетирования. С учетом опыта Принстонского университета вопрос о введении ограничений на долю отличных оценок не обсуждался.

В Корнельском университете проведен эксперимент по более широкому раскрытию информации о распределении оценок на разных дисциплинах. В 1996 г. началась публикация данных о медианных оценках по дисциплинам на сайте университета и скорректирован формат академической справки, в которую стали включать как оценки, выставленные студентам по дисциплинам, так и данные о средних (медианных) оценках. Целью этих мер было улучшить информированность студента и в конечном счете работодателя о реальных результатах обучения студента. Результатом реформы стал активный переток студентов на дисциплины, на которых выставлялись более высокие оценки, и усиление инфляции оценок. Решениями преподавательских собра-

⁵ Sharp rise in top degrees at three English universities investigated: <https://www.theguardian.com/education/2022/sep/02/office-for-students-investigates-sharp-rise-top-degrees-universities> (дата доступа 20.08.2024).

ний Корнельского университета все элементы реформирования были поэтапно отменены [Bar, Kadiyali, Zussman, 2009].

На рубеже 2020-х годов с устойчивым трендом завышения оценок столкнулась Высшая школа экономики⁶. Описывая эту ситуацию, проректор НИУ ВШЭ В.В. Радаев связал ее как с усилением давления со стороны студентов, заинтересованных в снижении платы за обучение, так и с изменением поведения преподавателей, мотивированных к повышению своего профессионального рейтинга и снижению издержек в переговорах со студентами [Радаев, 2023. С. 64–65]. Ситуация была скорректирована в результате введения административных требований по сокращению применения высших баллов при оценивании студентов и изменения условий контрактации преподавателей, практиковавших завышение оценок.

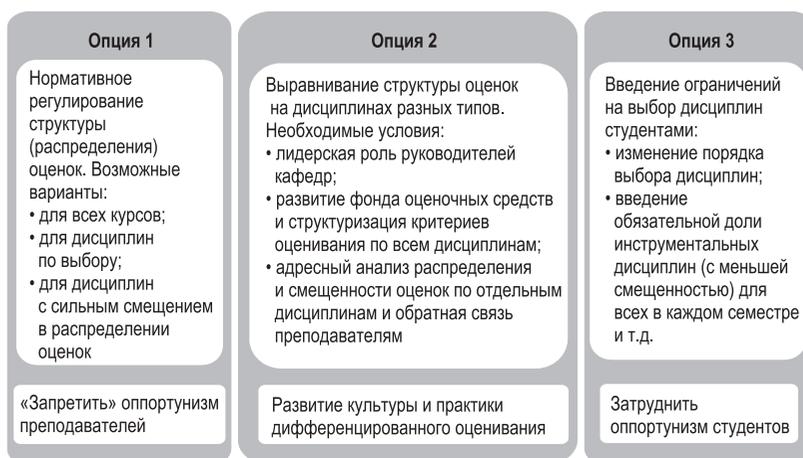
5.2. Стратегические альтернативы

На рис. 4 представлены три основных направления деятельности учебного заведения по изменению положения дел с инфляцией оценок. Они не являются взаимоисключающими, и каждый конкретный пакет мероприятий должен включать комбинацию шагов, предусматриваемых разными направлениями. Кроме того, очевидно, что усилия по противодействию инфляции оценок могут дать устойчивые результаты только при проведении институциональных и педагогических изменений, которые выходят далеко за рамки административного установления количественных пропорций в распределении оценок. При осуществлении мер, направленных на преодоление инфляции оценок, нужно учитывать мнение всех стейкхолдеров — студентов, преподавателей разных факультетов и кафедр, желательно добиться от них консенсуса в поддержке проводимой работы и обеспечить лидерство руководителей кафедр в работе по контролю над инфляцией оценок. Необходимы систематический сбор и анализ информации по распределению оценок в рамках промежуточного и итогового контроля знаний, на основе которой строятся мероприятия по изменению поведения студентов разного уровня подготовки при выборе дисциплин.

Количественные нормативы на распределение оценок или на долю отличных и хороших оценок (опция 1) могут вводиться для всех дисциплин или для дисциплин определенного вида. Применение нормативных ограничений должно сопровождаться изменением методик выставления оценок и расширением фонда оценочных средств, отработкой методик дифференцированной оценки знаний студентов. При кажущейся простоте такого подхода его применение может вызвать увеличение количества жа-

⁶ Комитет по качеству образования (2021) Инфляция оценок в НИУ ВШЭ: <https://www.hse.ru/mirror/pubs/share/474774074.pdf> (дата обращения: 20.08.2024).

Рис. 4. Способы снижения инфляции оценок



Источник: Составлено авторами.

лоб со стороны студентов и рост издержек на стороне преподавателей. При введении нормативов оценивания неизбежен также рост административных издержек контроля.

Вторая опция влияния на инфляцию оценок предполагает проведение системы методических мероприятий, которые должны обеспечить сохранение дифференциации оценок на дисциплинах разных направлений. Сегодня в дисциплинах с более высоким компонентом количественных, формализованных знаний смещенность в распределении оценок меньше, чем в дисциплинах, оперирующих преимущественно качественными категориями. Выбирая для изучения дисциплины, которые в наибольшей степени соответствуют их потенциальным способностям и уровню предшествующей подготовки, студенты могут выполнить учебную программу с меньшими (недостаточными) затратами усилий, и такая возможность искажает стимулы для студентов в выборе учебных дисциплин. Успешность реализации мероприятий по выравниванию структуры оценок на дисциплинах разных типов зависит от уровня методического обеспечения, полноты и качества оценочных средств по всем дисциплинам, активности преподавательских команд в обеспечении дифференциации оценок, лидерства руководителей кафедр и академических подразделений, осуществления мониторинга и обеспечения обратной связи с преподавателями.

Третье направление работы по снижению инфляции оценок составляют нормативные меры корректировки поведения студентов, с тем чтобы ограничить выбор ими дисциплин с высокой долей отличных оценок и сильным смещением в распределении оценок.

5.3. Система оценивания как общественный институт

Сложившаяся система оценивания и распределения оценок выстраивалась на протяжении длительного времени — годами и даже десятилетиями. Можно сказать, что система оценивания — это не только социальная практика, но даже некоторая социальная норма [Finefter-Rosenbluh, Levinson, 2015]. На уровне отдельных вузов правила и практики оценивания связаны с традициями определенных педагогических коллективов. Эти традиции отражают представления о справедливости и академической честности, на них могут влиять упрощенные понятия о равенстве и популистские идеи, которые создают этическую возможность для широкого распространения инфляции оценок.

Учитывая влияние на инфляцию оценок культурных и поведенческих факторов, на последнем этапе исследования мы вынесли полученные результаты на обсуждение с преподавателями. С целью разработки способов снижения смещенности в распределении оценок по дисциплинам преподавателям был предложен специально разработанный опросник. Одновременно обсуждались предложения по усовершенствованию балльно-рейтинговой системы оценивания.

Полученные в исследовании результаты имели большой отклик у преподавателей, что подтвердило актуальность и практическую значимость проведенной работы. В ходе обсуждения звучали предложения проводить регулярный мониторинг распределения оценок; обеспечить прозрачность информации о распределении оценок, выставляемых разными преподавателями; установить долю итоговой работы в формате экзамена на уровне 30% от суммы накопленных баллов; использовать единые критерии при проверке контрольных работ и экзаменов, проводить работы одновременно во всех группах; устные экзамены сдавать комиссиям из преподавателей.

Перспективными направлениями совершенствования системы оценивания преподаватели считают введение порогового уровня баллов для успешной сдачи экзамена; дифференцирование оценок по групповым проектам в зависимости от вклада каждого студента; обсуждение возможностей дифференциации системы оценивания в дисциплинах, требующих использования количественных методов, и регуляцию максимально разрешенной доли групповой работы по каждому предмету в дисциплинах социально-гуманитарного блока.

В целом проведенное обсуждение итогов исследования, как и обзор литературы, свидетельствует о том, что академическое сообщество и администраторы вузов озабочены весьма устойчивым трендом на инфляцию оценок. Идет поиск взаимоприемлемых подходов, направленных на сохранение традиций, с одной стороны, и на внедрение принципов дифференцированного оценивания знаний студентов — с другой. Уже разработано множество

подходов к работе с инфляцией оценок, и в каждом вузе требуется сформировать комплекс мер с учетом сложившихся традиций в преподавательских коллективах с дальнейшей проверкой его работоспособности на практике.

6. Заключение Значимость проблемы инфляции оценок возрастает по мере расширения охвата молодежи высшим образованием. Завершившийся в большинстве индустриальных стран переход от элитного к массовому высшему образованию предопределил высокую вероятность распространения феномена инфляции оценок. Проблема дополнительно осложняется с распространением платных форм обучения — и в связи с ожиданиями студентов, и в связи с возникновением искажающих стимулов на стороне администрации вузов, для которой весьма актуальной задачей становится сохранение контингента внебюджетных студентов.

Нетрудно привести целый список внешних ограничений, препятствующих решению проблемы инфляции оценок. Но сложность проблемы не снимает ответственности с сегодняшних администраторов систем образования и педагогического сообщества за предоставление качественного образования, которое предполагает сохранение системы дифференцированного оценивания на основе выполнения академических стандартов. Понимание феномена инфляции оценок и разработка инструментов его измерения и анализа является важным исходным основанием для того, чтобы успешно продвигаться к решению проблемы.

В проведенном исследовании явление инфляции оценок в системе высшего образования рассмотрено в статическом, динамическом и дифференцированном разрезе. В ходе статического анализа предложен аналог коэффициента асимметрии Пирсона для сравнения и ранжирования распределений оценок на разных дисциплинах и определения их смещенности относительно бенчмарка. Этот впервые примененный аналитический подход дает возможность более объективно выявлять инфляцию оценок на основе предложенной количественной характеристики. В качестве базового бенчмарка использован коллективный бенчмарк — среднее значение оценок по обязательным дисциплинам. Установлено, что оценки по элективным дисциплинам в среднем завышены относительно оценок по обязательным дисциплинам. Предложенный инструмент ранжирования может применяться преподавателями для сравнения результатов своего оценивания студентов с другими курсами или группами, что в итоге повысит дифференциацию оценок знаний и навыков учащихся.

В процессе статического анализа нами предложен подход к выявлению связи между завышением оценок и выбором студентами дисциплин, а также определены характеристики дисциплин,

в которых выставляются более высокие оценки: это более высокая доля групповых проектов, устный экзамен, меньшая доля контрольных работ, т.е. те характеристики, которые снижают возможность применения сопоставимых, четко прописанных критериев выставления баллов для каждого студента.

Предпосылкой к наличию динамической инфляции оценок на рассмотренном 10-летнем периоде не обнаружено. Также в ходе регрессионного анализа выявлено отсутствие дифференцированной инфляции оценок, которая предположительно могла бы выражаться в более лояльном оценивании в группах, изначально менее подготовленных к обучению.

Обсуждение полученных результатов с преподавательским сообществом позволило сформулировать возможные меры, направленные на снижение инфляции оценок. Предложенные мероприятия достаточно универсальны, что позволяет апробировать их в разных вузах с учетом специфики конкретного учебного заведения.

Благодарности Авторы выражают благодарность коллегам по экономическому факультету МГУ им. М.В. Ломоносова, членам Учебно-методической комиссии экономического факультета МГУ, организаторам и участникам семинара Института образования НИУ ВШЭ.

Литература

1. Абрамова М.О., Филькина А.В. (2023) Оценка студентами преподавания в университетах: больше вреда или пользы? *Высшее образование в России*, т. 32, № 8–9, сс. 130–146. <https://doi.org/10.31992/0869-3617-2023-32-8-9-130-146>
2. Ларионова В.А., Гончарова Н.В., Дайнеко Л.В. (2022) Проблемы инфляции оценок и симптомы академического мошенничества на МООК: о чем говорит учебная аналитика. *Университетское управление: практика и анализ*, т. 26, № 4, сс. 5–21. <http://doi.org/10.15826/umpra.2022.04.027>
3. Павлова Н.А. (2009) Повышение качества образования в высшей школе США: проблемы и направления развития. *Известия высших учебных заведений. Поволжский регион. Гуманитарные науки*, № 3, сс. 165–176.
4. Пискунов А.Ф., Филимонова Е.Н. (2006) Инфляция оценок: реальная или мнимая угроза? *Оценивание: образовательные возможности. Сб. научно-методических статей*. Минск: БГУ, вып. 4, сс. 78–87.
5. Радаев В.В. (2023) *Преподавание в кризисе*. М.: НИУ ВШЭ.
6. Achen A.C., Courant P.N. (2009) What Are Grades Made of? *Journal of Economic Perspectives*, vol. 23, no 3, pp. 77–92. <http://dx.doi.org/10.1257/jep.23.3.77>
7. Bar T., Kadiyali V., Zussman A. (2009) Grade Information and Grade Inflation: The Cornell Experiment. *Journal of Economic Perspectives*, vol. 23, no 3, pp. 93–108. <http://dx.doi.org/10.1257/jep.23.3.93>
8. Chan W., Hao L., Suen W. (2007) A Signaling Theory of Grade Inflation. *International Economic Review*, vol. 48, no 3, pp. 1065–1090. <http://dx.doi.org/10.1111/j.1468-2354.2007.00454.x>
9. Chowdhury F. (2018) Grade Inflation: Causes, Consequences and Cure. *Journal of Education and Learning*, vol. 7, no 6, pp. 86–92. <http://dx.doi.org/10.5539/jel.v7n6p86>

10. Eiszler C.F. (2002) College Students' Evaluations of Teaching and Grade Inflation. *Research in Higher Education*, vol. 43, no 4, pp. 483–501. <http://dx.doi.org/10.1023/A:1015579817194>
11. Finefter-Rosenbluh I., Levinson M. (2015) What Is Wrong with Grade Inflation (if Anything)? *Philosophical Inquiry in Education*, vol. 23, no 1, pp. 3–21. <http://dx.doi.org/10.7202/1070362ar>
12. Gershenson S. (2018) *Grade Inflation in High Schools (2005–2016)*. Available at: https://fordhaminstitute.org/sites/default/files/20180919-grade-inflation-high-schools-2005-2016_0.pdf (accessed 20 August 2024).
13. Johnson V.E. (2006) *Grade Inflation: A Crisis in College Education*. New York, NY: Springer Science & Business Media.
14. Juola A.E. (1976) *Grade Inflation in Higher Education: What Can or Should We Do?* Paper presented at the Annual Meeting of National Council on Measurement in Education (San Francisco, CA, April 1979). Available at: <https://goo.su/fbXa> (accessed 20 August 2024).
15. Office for Students (2022) *Analysis of Degree Classifications over Time: Changes in Graduate Attainment from 2010–11 to 2021–22*. Available at: <https://www.officeforstudents.org.uk/publications/analysis-of-degree-classifications-over-time-changes-in-graduate-attainment-from-2010-11-to-2021-22/> (accessed 20 August 2024).
16. Sabot R., Wakeman-Linn J. (1991) Grade Inflation and Course Choice. *Journal of Economic Perspectives*, vol. 5, no 1, pp. 159–170. <http://dx.doi.org/10.1257/jep.5.1.159>
17. Stroebe W. (2020) Student Evaluations of Teaching Encourages Poor Teaching and Contributes to Grade Inflation: A Theoretical and Empirical Analysis. *Basic and Applied Social Psychology*, vol. 42, no 4, pp. 276–294. <http://dx.doi.org/10.1080/01973533.2020.1756817>
18. Tyner A., Gershenson S. (2020) Conceptualizing Grade Inflation. *Economics of Education Review*, vol. 78, October, Article no 102037. <http://dx.doi.org/10.1016/j.econedurev.2020.102037>

References

- Abramova M.O., Filkina A.V. (2023) Student Evaluation of Teaching: More Harm than Good? *Vysshee obrazovanie v Rossii / Higher Education in Russia*, vol. 32, no 8–9, pp. 130–146 (In Russian). <https://doi.org/10.31992/0869-3617-2023-32-8-9-130-146>
- Achen A.C., Courant P.N. (2009) What Are Grades Made of? *Journal of Economic Perspectives*, vol. 23, no 3, pp. 77–92. <http://dx.doi.org/10.1257/jep.23.3.77>
- Bar T., Kadiyali V., Zussman A. (2009) Grade Information and Grade Inflation: The Cornell Experiment. *Journal of Economic Perspectives*, vol. 23, no 3, pp. 93–108. <http://dx.doi.org/10.1257/jep.23.3.93>
- Chan W., Hao L., Suen W. (2007) A Signaling Theory of Grade Inflation. *International Economic Review*, vol. 48, no 3, pp. 1065–1090. <http://dx.doi.org/10.1111/j.1468-2354.2007.00454.x>
- Chowdhury F. (2018) Grade Inflation: Causes, Consequences and Cure. *Journal of Education and Learning*, vol. 7, no 6, pp. 86–92. <http://dx.doi.org/10.5539/jel.v7n6p86>
- Eiszler C.F. (2002) College Students' Evaluations of Teaching and Grade Inflation. *Research in Higher Education*, vol. 43, no 4, pp. 483–501. <http://dx.doi.org/10.1023/A:1015579817194>
- Finefter-Rosenbluh I., Levinson M. (2015) What Is Wrong with Grade Inflation (if Anything)? *Philosophical Inquiry in Education*, vol. 23, no 1, pp. 3–21. <http://dx.doi.org/10.7202/1070362ar>
- Gershenson S. (2018) *Grade Inflation in High Schools (2005–2016)*. Available at: https://fordhaminstitute.org/sites/default/files/20180919-grade-inflation-high-schools-2005-2016_0.pdf (accessed 20 August 2024).

- Johnson V.E. (2006) *Grade Inflation: A Crisis in College Education*. New York, NY: Springer Science & Business Media.
- Juola A.E. (1976) *Grade Inflation in Higher Education: What Can or Should We Do?* Paper presented at the Annual Meeting of National Council on Measurement in Education (San Francisco, CA, April 1979). Available at: <https://goo.su/fbXa> (accessed 20 August 2024).
- Larionova V.A., Goncharova N.V., Daineko L.V. (2022) Problems of Grade Inflation and Symptoms of Academic Fraud within the MOOCs: What Educational Analytics Says. *University Management: Practice and Analysis*, vol. 26, no 4, pp. 5–21. <http://doi.org/10.15826/umpa.2022.04.027> (In Russian).
- Office for Students (2022) *Analysis of Degree Classifications over Time: Changes in Graduate Attainment from 2010–11 to 2021–22*. Available at: <https://www.office-forstudents.org.uk/publications/analysis-of-degree-classifications-over-time-changes-in-graduate-attainment-from-2010-11-to-2021-22/> (accessed 20 August 2024).
- Pavlova N.A. (2009) Improving the Quality of Education in Higher Education in the USA: Problems and Directions of Development. *University Proceedings. Volga region. Humanities*, no 3, pp. 165–176 (In Russian).
- Piskunov A.F., Filimonova E.N. (2006) Inflation of Estimates: A Real or Imaginary Threat? *Assessment: Educational Opportunities. Collection of Scientific and Methodological Articles*, Minsk: BSU, iss. 4, pp. 78–87 (In Russian).
- Radaev V.V. (2023) *Teaching in Crisis*. Moscow: HSE (In Russian).
- Sabot R., Wakeman-Linn J. (1991) Grade Inflation and Course Choice. *Journal of Economic Perspectives*, vol. 5, no 1, pp. 159–170. <http://dx.doi.org/10.1257/jep.5.1.159>
- Stroebe W. (2020) Student Evaluations of Teaching Encourages Poor Teaching and Contributes to Grade Inflation: A Theoretical and Empirical Analysis. *Basic and Applied Social Psychology*, vol. 42, no 4, pp. 276–294. <http://dx.doi.org/10.1080/01973533.2020.1756817>
- Tyner A., Gershenson S. (2020) Conceptualizing Grade Inflation. *Economics of Education Review*, vol. 78, October, Article no 102037. <http://dx.doi.org/10.1016/j.econedurev.2020.102037>