

# ПРЕДСКАЗАНИЕ УСПЕВАЕМОСТИ В НИУ ВШЭ ПО РЕЗУЛЬТАТАМ ЕГЭ

---

## Введение

С 2009 г. в Российской Федерации при приеме в государственные высшие учебные заведения в качестве вступительных испытаний повсеместно используются результаты ЕГЭ. Студенты, поступившие в 2009 г., закончили свой первый курс в 2010 г. Появилась информационная база для оценивания того, насколько результаты ЕГЭ абитуриентов позволяют предсказать их последующую успеваемость в вузе. Исследование подобной связи важно с точки зрения оценивания валидности ЕГЭ как критерия отбора в вузы, а также дает статистическую основу для выработки рекомендаций по улучшению критериев отбора.

Если успеваемость в вузе и результаты ЕГЭ статистически слабо связаны, то закономерно возникает вопрос о том, насколько оправданно использование ЕГЭ для отбора в вузы на места, финансируемые из государственного бюджета. Если ЕГЭ в целом оказывается адекватным предиктором вузовской успеваемости, то его использование в целом оправданно, но возможно рассматривать варианты более эффективного использования информации о баллах ЕГЭ абитуриентов при приеме.

В настоящее время критерием, по которому осуществляется ранжирование абитуриентов, является равновзвешенная сумма баллов, полученных на ЕГЭ по избранному предметам. Перечень предметов из трех или четырех вступительных испытаний по каждому образовательному направлению определяется Министерством образования и науки Российской Федерации. В это число обязательно входят экзамены по русскому языку и профильному общеобразовательному предмету.

Выбор дисциплин ЕГЭ и весовых коэффициентов при них при суммировании баллов играет важную роль. Различные дисциплины единого экзамена в некоторой степени тестируют различные специфические способности и знания

учащихся. Например, для направления «Физика» оптимально оценивать знания по физике, для направления «Химия» – по химии и т.п. Весовые коэффициенты определяют вклад отдельных дисциплин в итоговый критерий. Не очевидно, что использование равных весов, предусмотренное текущими правилами приема, является лучшим методом.

При неоптимальном выборе дисциплин вступительных испытаний и их весов происходит неблагоприятный отбор абитуриентов. В вуз могут попасть учащиеся, средняя успеваемость которых ниже, чем ожидаемая успеваемость некоторых абитуриентов, отсеянных во время приемной кампании. Выбор оптимального набора дисциплин и их весов может быть основан на регрессионном анализе модели, в которой в качестве объясняемой переменной выступает вузовская успеваемость, а в роли объясняющих факторов – результаты вступительных испытаний.

Использование стандартизированных тестов при приеме в вузы является новой практикой для России. За рубежом, особенно в США, накоплен значительный опыт в области изучения связи результатов довузовского тестирования и последующей успеваемости в вузе. В США многие вузы учитывают при приеме результаты стандартизированных тестов, таких как SAT или ACT, проводимых частными организациями в течение многих десятилетий. Статья [Burton, Ramist, 2001] содержит обзор работ 1980–2000 гг. по изучению влияния результатов SAT на академическую успеваемость в вузах. Работа [Kobrin et al., 2008] базируется на данных о 150 тыс. учащихся, принятых по SAT в 110 колледжей и университетов. Большинство исследований подтверждают статистически значимую связь результатов довузовского тестирования и последующей успеваемости в вузе.

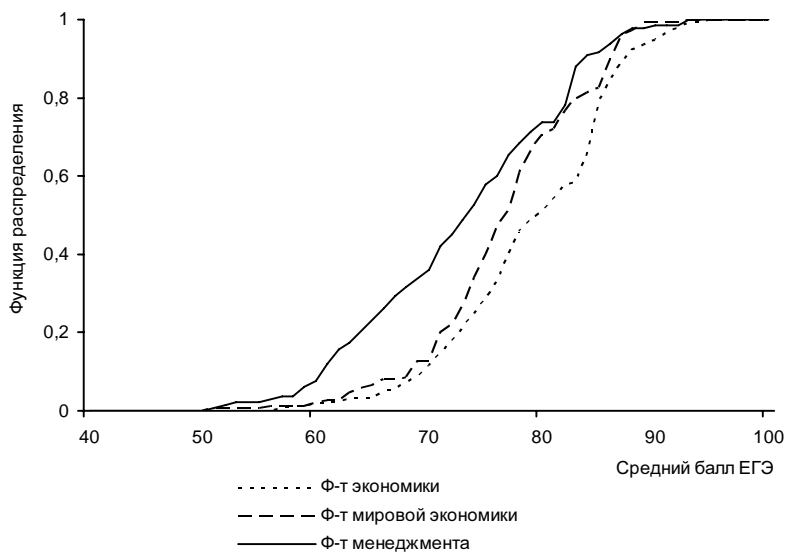
В статье [Польдин, 2011] исследовались зависимости среднего балла после первого курса от результатов вступительных испытаний для студентов факультета экономики НИУ ВШЭ. В настоящей работе связь довузовской и вузовской успеваемости исследуется по значительно большему числу студентов, обучающихся на крупнейших факультетах ВШЭ – экономики, мировой экономики и мировой политики, менеджмента.

## **Сравнение успеваемости поступивших по результатам олимпиад и ЕГЭ**

По правилам приема в вузы в 2009 г. без вступительных испытаний зачислялись победители и призеры некоторых олимпиад. На оставшиеся бюд-

жетные места осуществлялось зачисление по конкурсу баллов ЕГЭ. При поступлении на факультеты экономики, мировой экономики и политики, менеджмента ВШЭ учитывались результаты ЕГЭ по четырем дисциплинам: математике, обществознанию, русскому языку, иностранному языку. Часть не прошедших по конкурсу абитуриентов были зачислены на места с оплатой стоимости обучения на договорной основе.

На рис. 1 приведены эмпирические функции распределения среднего балла ЕГЭ студентов трех факультетов, зачисленных в 2009 г. по результатам ЕГЭ (как на бюджетные места, так и по договорам) и закончивших первый курс в 2010 г. Таких студентов оказалось 157 человек для факультета экономики, 150 человек – для факультета мировой экономики, 133 человека – для факультета менеджмента. Данные о результатах ЕГЭ абитуриентов ВШЭ находились в открытом доступе на сайте ВШЭ ([www.hse.ru](http://www.hse.ru)).

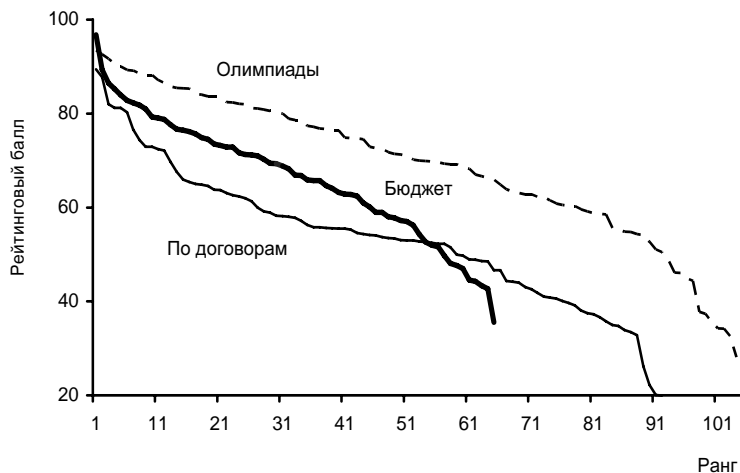


**Рис. 1.** Эмпирические интегральные функции распределения среднего балла ЕГЭ: факультет экономики, факультет мировой экономики, факультет менеджмента

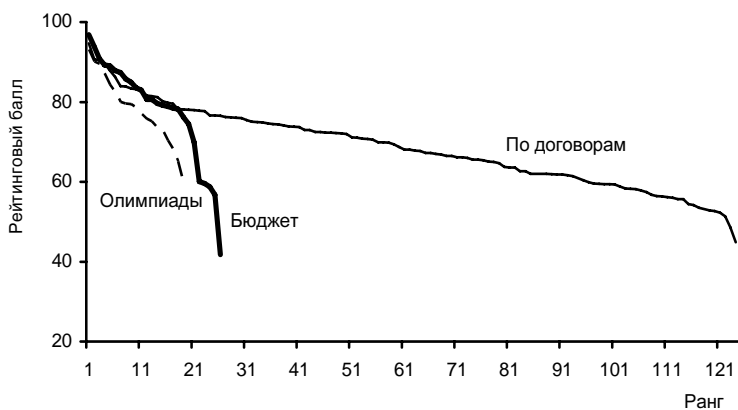
Показателем, характеризующим общую успеваемость студента ВШЭ, является рейтинг – взвешенная сумма оценок по дисциплинам учебной программы, весами служат кредиты учебной нагрузки по предмету. Текущие рейтинги студентов за предыдущий семестр и их кумулятивные рейтинги за весь период

обучения размещаются на сайте ВШЭ. В настоящей работе для удобства рейтинговые баллы нормированы так, что их максимально возможное значение составляет 100.

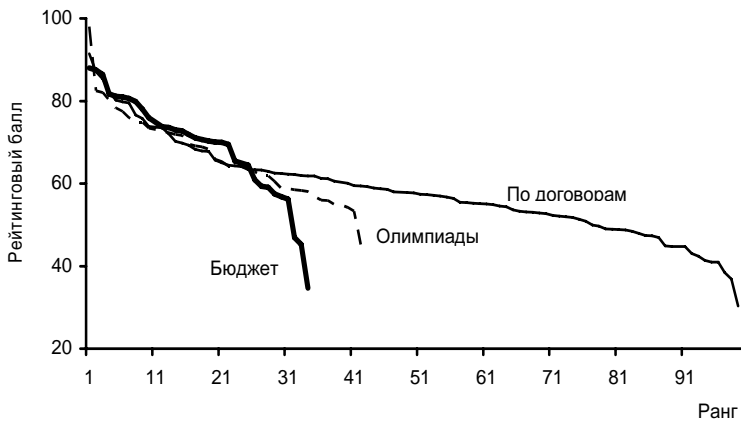
На рис. 2 приведены ранжированные в убывающем порядке рейтинговые баллы после первого года обучения для студентов, поступивших по результатам ЕГЭ на бюджетные места, договорные (платные) места, а также зачисленных без вступительных испытаний в качестве победителей и призеров олимпиад.



а) Факультет экономики



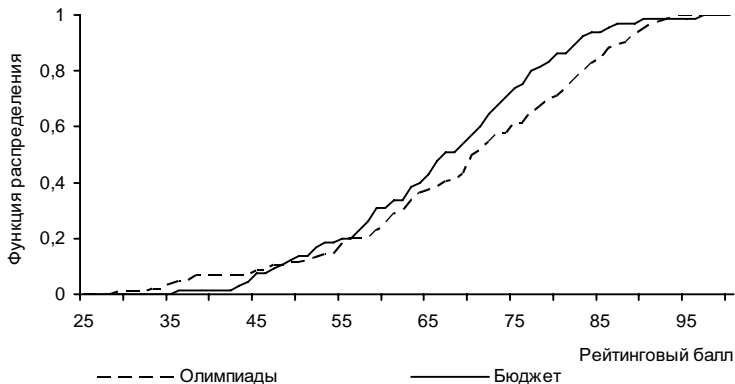
б) Факультет мировой экономики



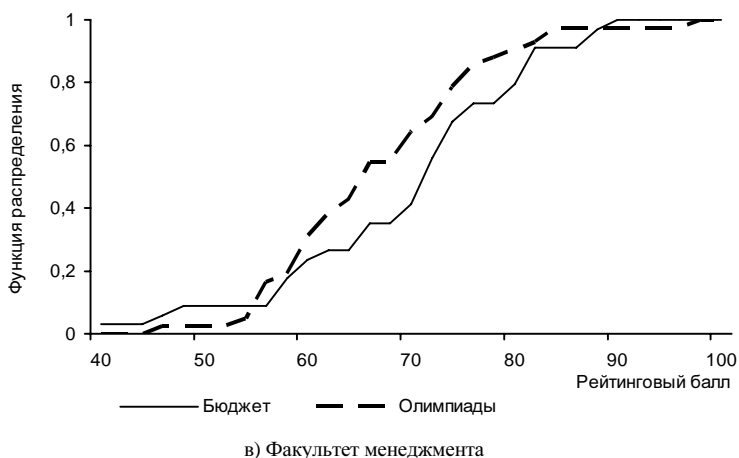
в) Факультет менеджмента

**Рис. 2.** Ранжированные рейтинги студентов по итогам первого года

Представляет интерес сравнение успеваемости студентов, поступивших по конкурсу баллов ЕГЭ, и студентов, зачисленных вне конкурса в качестве победителей и призеров олимпиад. Для этого целесообразно сравнивать их эмпирические интегральные функции распределения (см. рис. 3). Видно, что успеваемость этих двух категорий студентов довольно близка. Для факультета экономики среди наиболее успешных студентов с высоким рейтингом несколько преобладают олимпиадники, на двух других факультетах ситуация обратная. Это можно объяснить тем, что на факультете экономики велика доля победителей и призеров Всероссийской олимпиады школьников.



а) Факультет экономики



**Рис. 3.** Эмпирические интегральные функции распределения рейтингов студентов

## Связь ЕГЭ и рейтинга студентов после первого курса

Для анализа связи среднего балла ЕГЭ при поступлении и рейтинга студентов после первого курса рассматривались две спецификации уравнений регрессии. В первой модели в качестве регрессора выступает средний балл ЕГЭ, усредненный по четырем дисциплинам вступительных испытаний. Именно этот показатель использовался как критерий ранжирования при конкурсном отборе.

Во второй спецификации регрессорами выступали баллы по отдельным предметам. В обеих спецификациях зависимой переменной является рейтинговый балл по итогам первого года обучения.

В качестве методов оценивания применялись медианная регрессия и робастная регрессия, оценки которых менее чувствительны к большим возмущениям по сравнению с оценками метода наименьших квадратов. В медианной регрессии минимизируется сумма абсолютных значений остатков регрессии:

$$\hat{\theta}_{L_1} = \arg \min_{\theta} \sum_{i=1}^n |r_i(\theta)|, \quad r_i(\theta) = y_i - X\theta.$$

В робастной регрессии вместо модуля остатков использовалась другая функция потерь ( $\rho$ ):

$$\hat{\theta}_M = \arg \min_{\theta} \sum_{i=1}^n \rho \left( \frac{r_i(\theta)}{\sigma} \right),$$

$$\rho(u) = \begin{cases} 1 - \left[ 1 - \left( \frac{u}{k} \right)^2 \right]^3, & \text{если } |u| \leq k, \\ 1, & \text{если } |u| > k; \end{cases} \quad \text{где } k = 4,685.$$

В табл. 1 представлены оценки коэффициентов парной регрессии. Коэффициент при среднем балле ЕГЭ характеризует прирост ожидаемого рейтинга в вузе при увеличении среднего балла ЕГЭ на единицу. Наибольшие значения этого показателя оказались на факультете экономики, наименьшее – на факультете менеджмента.

Коэффициент детерминации  $R^2$  является распространенной характеристикой объясняющей силы регрессии. Он измеряет долю вариации рейтинга, объясняемую вариацией баллов ЕГЭ. Эта величина составила 32–34%. Для медианной регрессии аналогичным показателем является псевдо- $R^2$ , его величина составила 18–21%.

Механически сравнивать обладателей одинаковых рейтингов, обучающихся на различных факультетах, было бы не вполне корректно. Но нельзя сказать, что подобное сравнение невозможно. Например, для сопоставления успеваемости можно учитывать полученные регрессионные зависимости. На рис. 4 построены линии регрессии для оценок, рассчитанных методом робастной регрессии. Зависимости имеют смысл ожидаемых значений рейтинговых баллов, предсказанных по среднему баллу ЕГЭ. Труднее всего получить высокие рейтинги на факультете экономики, наиболее мягкая система оценок оказалась на факультете мировой экономики.

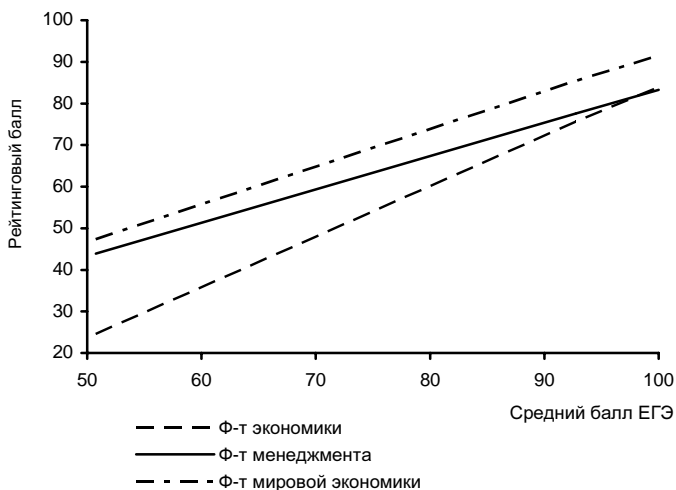
**Таблица 1.**

Связь рейтингового балла в вузе и среднего балла ЕГЭ, робастная регрессия (1) и медианная регрессия (2)

Регрессоры	Факультет экономики		Факультет мировой экономики		Факультет менеджмента	
	1	2	1	2	1	2
Константа	-37,098 (10,812)	-25,743 (14,440)	1,413 (8,267)	2,902 (10,036)	3,210 (7,163)	6,236 (7,646)
Средний балл ЕГЭ	1,215*** (0,136)	1,074*** (0,182)	0,906*** (0,107)	0,887*** (0,130)	0,802*** (0,097)	0,768*** (0,104)
R <sup>2</sup>	0,340		0,325		0,341	
Псевдо-R <sup>2</sup>		0,208		0,210		0,182

Примечания. В скобках указаны стандартные погрешности оценок.

\*, \*\*, \*\*\* обозначают значимость на 10-процентном, 5-процентном и 1-процентном уровнях соответственно.



**Рис. 4.** Предсказанные по ЕГЭ вузовские рейтинги

В табл. 2 приведены оценки коэффициентов множественной регрессии, в которой в качестве объясняющих переменных выступают баллы ЕГЭ по отдельным дисциплинам. Из-за большего числа регрессоров качество предсказания в этой модели выше, чем в парной регрессии. Коэффициенты при дисциплинах имеют смысл приращений ожидаемого значения среднего балла в



вузе при увеличении балла ЕГЭ по данному предмету на единицу. Чем выше коэффициент, тем ценнее баллы по соответствующей дисциплине. На факультетах экономики и мировой экономики наибольший коэффициент имеет математика, на факультете менеджмента – обществознание. Иностранный язык оказался незначим на факультете экономики, а на факультете менеджмента – математика и русский язык.

**Таблица 2.**

Связь рейтингового балла в вузе и баллов по предметам ЕГЭ, робастная регрессия (1) и медианная регрессия (2)

Регрессоры	Факультет экономики		Факультет мировой экономики		Факультет менеджмента	
	1	2	1	2	1	2
Константа	-43,804 (10,906)	-32,578 (14,379)	1,251 (8,867)	4,892 (10,544)	-3,361 (8,456)	0,158 (10,867)
Математика	0,624*** (0,129)	0,660*** (0,169)	0,341*** (0,102)	0,340*** (0,121)	0,153 (0,109)	0,134 (0,141)
Русский язык	0,250* (0,120)	0,106 (0,165)	0,202* (0,107)	0,182 (0,122)	0,198 (0,127)	0,196 (0,165)
Обществознание	0,314* (0,133)	0,306* (0,181)	0,212** (0,091)	0,150 (0,106)	0,376** (0,156)	0,344* (0,202)
Иностранный язык	0,144 (0,112)	0,127 (0,153)	0,169** (0,069)	0,200** (0,079)	0,170** (0,070)	0,177* (0,093)
R <sup>2</sup>	0,376		0,345		0,361	
Псевдо-R <sup>2</sup>		0,237		0,214		0,187

*Примечания.* В скобках указаны стандартные погрешности оценок.

\*, \*\*, \*\*\* обозначают значимость на 10-процентном, 5-процентном и 1-процентном уровнях соответственно.

Если предположить, что оцененные регрессии достоверно характеризуют зависимости между переменными в генеральной совокупности, то на основе полученных регрессионных коэффициентов при дисциплинах ЕГЭ возможно получить весовые коэффициенты при этих дисциплинах ЕГЭ для формирования критерия отбора, дающего оптимальный прогноз успеваемости. Для этого каждый коэффициент при ЕГЭ из табл. 2 (константа не учитывается) нужно разделить на сумму всех коэффициентов. Полученные веса в процентах приведены в табл. 3.

Расчитанные по регрессиям оптимальные коэффициенты отличаются от фактически используемых 25-процентных весов для каждого предмета, вслед-

ствие чего возникают искажения при отборе студентов, зачисляемых в вуз на бюджетные места. Например, на факультете экономики оптимальный вес математики примерно равен весу остальных трех гуманитарных дисциплин вместе взятых. Но по правилам приема вес математики составляет только 25%. В результате абитуриенты с высоким средним баллом, полученным за счет гуманитарных дисциплин, вытеснили кандидатов, более одаренных математически, хотя последние имели лучшие перспективы для обучения.

**Таблица 3.** Весовые коэффициенты при дисциплинах ЕГЭ, рассчитанные по робастной регрессии (1) и медианной регрессии (2), %

	Факультет экономики		Факультет мировой экономики		Факультет менеджмента	
	1	2	1	2	1	2
Математика	47	55	37	39	17	16
Русский язык	19	9	22	21	22	23
Обществознание	24	26	23	17	42	40
Иностранный язык	11	11	18	23	19	21

## Заключение

В работе исследовались статистические связи успеваемости студентов на первом курсе с результатами ЕГЭ при поступлении для трех факультетов ВШЭ. Расчеты проводились по данным приема 2009 г. Получены оценки коэффициентов регрессий, в которых в качестве независимых переменных выступали средний балл ЕГЭ по четырем дисциплинам ЕГЭ и баллы по отдельным предметам ЕГЭ. Для оценивания использовались медианная регрессия и робастная регрессия как более устойчивые к выбросам по сравнению с методом наименьших квадратов.

Средний балл ЕГЭ по математике, русскому и иностранному языкам, обществознанию способен значимо предсказывать успеваемость в вузе по итогам первого года. В регрессиях с использованием результатов отдельных экзаменов удается достичь несколько лучшего качества предсказания. Вклады ЕГЭ по различным дисциплинам в объяснение вузовского рейтинга различаются. Эта роль варьируется в зависимости от программы обучения. На факультетах экономики и мировой экономики наибольшее влияние на рейтинг имеет ЕГЭ по математике, на факультете менеджмента – ЕГЭ по обществознанию.

## Литература

*Польдин О.В.* Прогнозирование успеваемости в вузе по результатам ЕГЭ // Прикладная эконометрика. 2011. № 1. С. 56–69.

*Burton N.W., Ramist L.* Predicting Success in College: SAT Studies of Classes Graduating Since 1980 // College Board Research Report. № 2001-2. N.Y.: The College Board, 2001.

*Kobrin J.L., Patterson B.F., Shaw E.J. et al.* Validity of the SAT for Predicting First-Year College Grade Point Average // College Board Research Report. № 2008-5. N.Y.: The College Board, 2008.