

Е.В. Балацкий

ЦЭМИ РАН, Москва

Н.А. Екимова

Государственный университет управления, Москва

Сравнительная надежность глобальных рейтингов университетов¹

В статье рассматривается проблема множественности глобальных рейтингов университетов, дана их классификация. Осуществлена оценка взаимной согласованности рейтингов университетов. Предложена и апробирована процедура оценки и сравнения уровня надежности существующих глобальных рейтингов.

Ключевые слова: *рейтинги университетов, критерии ранжирования, образование, экономическая политика, локальная надежность, глобальная надежность.*

Классификация JEL: I21, I23.

1. Введение

С 2003 г. на академический мир обрушилась лавина глобальных рейтингов университетов, которые с каждым годом усиливают свое влияние на разные стороны жизни общества. Апофеозом этого влияния может служить голландский иммиграционный закон, согласно которому приоритет на въезд в страну предоставляется людям с высокой профессиональной квалификацией, которую подтверждает диплом университета из списка Топ-150 глобального рейтинга² (Batu, 2010). Одновременно с этим многими странами осуществляются многомиллиардные инициативы, направленные на создание университетов мирового класса. Между тем набирает остроту проблема адекватности рейтинговых оценок. Эффект аберрации истинного положения дел усугубляется гипертрофированным влиянием рейтингов университетов на реальную жизнь общества.

К настоящему моменту накоплено уже достаточно много обзорных и критических работ, посвященных проблеме корректности рейтингов университетов. Один из последних обзоров текущих глобальных рейтингов университетов дан в (Балацкий, Екимова, 2011). Проблемы манипулирования рейтинговыми данными нашли свое отражение в работах (Batu, 2010; Diamond, Graham, 2000; Holmes, 2010; Usher, Savino, 2007; Van Raan, 2005; Smith, 2010; Marginson, 2007). Серьезные критические замечания в адрес библиометрических индексов были выдвинуты в (Arnold, Fowler, 2011). Дополнительные критические аргументы были приведены в работах (Свердлов, 2006; Арский, Маркусова, Чумакова, 2007; Гинзбург, 1998).

Множество содержательных проблем развития университетов в контексте формальных процедур формирования рейтингов были

¹ Статья подготовлена при финансовой поддержке Российского фонда фундаментальных исследований (проект 11-06-00210-а). Авторы выражают благодарность В.М. Полтеровичу, а также анонимному рецензенту за сделанные ими замечания.

² Для этого в Голландии существует собственный глобальный рейтинг университетов Leiden Ranking (LR), его составляет Лейденский университет (Centre for Science and Technology Studies, Leiden University). Такая работа для вузов Европы ведется с 2000 г., а для общемировой выборки – с 2003 г. Данный рейтинг основан на библиометрических индексах (The Leiden Ranking, 2008).

раскрыты в (Салми, 2009). Наличие же множества различных глобальных рейтингов привело к попыткам оценки степени их согласованности (Aguillo, Bar-Ilan, Levene, Ortega, 2010).

Однако новые рейтинги продолжают появляться. Так, в 2011 г. появился World Reputation Rankings (WRR), разработанный «Times Higher Education» совместно с агентством «Thomson Reuters» в качестве дополнения к ставшему уже популярным рейтингу ТНЕ и вместе с тем некоей альтернативы существующим рейтингам. Его особенность состоит в том, что уровень академической репутации вузов оценивается на основании данных, полученных в ходе опроса свыше 13 тыс. представителей международного академического сообщества из 131 страны. На основании результатов наукометрического анализа продуктивности и цитируемости различных авторов компанией «Thomson Reuters» были отобраны преподаватели, работающие в вузах более 16 лет и опубликовавшие более 50 научных работ, которым были разосланы анкеты на 8 языках (английском, немецком, французском, испанском, японском, китайском и двух португальских – европейском и бразильском). Эксперты должны были выбрать из сотен дисциплин и 6 тыс. вузов только 10 лучших, по их мнению, университетов, а также сильнейшие вузы для продолжения образования по магистерской и докторской программам. Основными оценочными критериями были международная студенческая мобильность, обеспечиваемая вузом, число международных стипендиальных программ и международных исследований. Результирующий показатель представляет собой средневзвешенную оценку образования с весом 1/3 и оценки исследований с весом 2/3. Все оценки нормированы по максимуму и приведены к 100-балльной шкале (The Times Higher Education, 2010a).

Таким образом, рейтинговое движение, изначально направленное на структурирование отрасли высшего образования, на нынешнем этапе развития порождает *метааналитическую проблему* – упорядочение самих глобальных рейтингов. Сегодня возникают вопросы, на которые пока нет адекватных ответов. Например, что делать университету, чтобы занять достойное место в глобальном рейтинге университетов? Как следует относиться к существующим рейтингам? Каким из них можно доверять, а каким – нельзя?

В связи с этим в данной статье рассматривается три группы вопросов. Первая предполагает классификацию существующих глобальных рейтингов, которая позволяет объективно позиционировать их достоинства и недостатки. Вторая группа вопросов предполагает оценку степени согласованности наиболее авторитетных глобальных рейтингов. И третья группа связана с определением наиболее достоверных рейтингов на базе разработанной для этого простой методики.

2. Классификация рейтингов

Обобщив разнообразные глобальные рейтинги университетов, дадим им достаточно простую, состоящую из четырех групп, классификацию.

Первую группу составляют *традиционные* рейтинги, которые основаны на оценке большого числа *объективных* показателей деятельности университетов (индексы активности научных публикаций и цитирований, показатели ресурсного обеспечения, успешности карьеры выпускников и пр.). К их числу относятся китайский Academic Ranking of World Universities (ARWU), тайваньский Performance Ranking of Scientific Papers for World Universities (PRSP), французский Professional Ranking of World Universities (PRWU) и голландский Leiden Ranking (LR) рейтинги.

Вторую группу образуют *репутационные*, или *персоналифицированные*, рейтинги, основанные на обработке мнений *экспертов*, полученных в результате специальных опросов. К числу таковых относится пока только британский рейтинг World Reputation Rankings (WRR).

Третью группу составляют *виртуальные*, или Интернет-рейтинги, которые основаны на показателях популярности университетских сайтов в Интернет-среде (частота посещений, число ссылок и пр.). К разряду таковых относятся: испанский Webometrics (Web), армянский «АркаЛер» и анонимный рейтинг World Universities Web Ranking (4ICU).

Четвертую группу образуют *смешанные*, или *гибридные*, рейтинги, при составлении которых используют «смесь» трех предыдущих методологий. В настоящее время – это, как правило, синтез традиционных и репутационных рейтингов. К их числу относятся английский Times Higher Education (THE), английский Quacquarelli Symonds (QS), американский Newsweek «The Top 100 Global Universities» (NTGU), американский U.S. News & World Report's World's Best Universities ranking (WBU) и российский Global Universities Ranking (GUR).

Распределение глобальных рейтингов по разным типам сведено в табл. 1.

Смысл предложенной классификации состоит в том, чтобы понять, существует ли связь между спецификой рейтингов и их надежностью. При наличии такой связи можно «изнутри» проанализировать источники возникающих ошибок в присвоении ранга тому или иному университету.

Таблица 1

Распределение глобальных рейтингов университетов по типам

Тип глобального рейтинга университетов			
Традиционный	Репутационный	Виртуальный	Гибридный
ARWU		Web	THE, QS
PRSP, LR	WRR	4ICU	NTGR, WBU
PRWU		АркаЛер	GUR

3. Симпатии и антипатии, альянсы и конкуренты

Проблема адекватности и надежности рейтинговых оценок является острой, однако в условиях множественности глобальных рейтингов университетов она дополнительно обостряется. Какой из существующих рейтингов следует признать лучшим? На какой из них следует ориентироваться? Какой рейтинг вызывает наибольшее доверие?

Сегодня все эти вопросы остаются без ответов. Между тем независимо от ответов на эти вопросы общественное мнение стихийно дрейфует в направлении некоторой группы глобальных рейтингов. Так, например, опрос, проведенный компанией «Thomson Reuters» совместно с «Ipsos Insight Corporation», ведущей исследовательской фирмой, занятой разработкой и внедрением онлайн-опросов, охватил 350 респондентов (студентов и сотрудников вузов) более чем из 30 стран мира (Smith, 2010). В частности, была определена популярность трех рейтинговых систем (табл. 2).

Результаты позволяют сделать следующие выводы.

Во-первых, наиболее популярным и узнаваемым глобальным рейтингом университетов является британский THE; на втором месте с большим отрывом стоит китайский ARWU, замыкает же опросный лист американский WBU (U.S. News & World Report, 2011).

Во-вторых, выявленная закономерность распространяется и на представителей Азии, которые, как это ни странно, отдают явное предпочтение европейскому рейтингу по сравнению с азиатским. Более того, жители Азии демонстрируют даже большую осведомленность о THE, чем представители всех других регионов.

В-третьих, американский рейтинг WBU, будучи формально глобальным, по сфере применения является все-таки скорее локальным и региональным; представители других частей света очень мало с ним знакомы.

Таким образом, можно констатировать, что между глобальными рейтингами на мировой арене развернулась борьба за призна-

Таблица 2

Уровень узнаваемости (знакомства) респондентов с разными рейтинговыми системами, 2009 г., %

Регион мира	Разновидность глобального рейтинга		
	THE	ARWU	WBU
Северная Америка	54	34	95
Европа	78	41	16
Австралия и Новая Зеландия	86	81	32
Азия	91	66	31
Всего	72	49	44

Источник: (Smith, 2010).

ние широких масс, которое закрепляется в результатах авторитетных опросов. При этом 85% респондентов считает сравнение академических рейтингов полезным. Следовательно, население желает иметь информацию о сильных и слабых сторонах существующих глобальных рейтингов.

Помимо прочего рейтинговые системы подвержены распаду и объединениям. Классическим примером первого процесса может служить рейтинг QS-THE, который с 2009 г. распался на два самостоятельных глобальных рейтинга – QS и THE. Сегодня это – две наиболее авторитетные рейтинговые системы. Параллельно оператор рейтинга THE «Times Higher Education» стал сотрудничать с агентством «Thomson Reuters», которое с этого момента стало заниматься не только информационной базой научного цитирования Web of Science, но и разработкой глобального рейтинга университетов. Таким образом, мы наблюдаем интеграцию двух информационных продуктов у одного разработчика, представленного альянсом двух фирм.

Однако объединения рейтингов часто принимают форму неявных альянсов. Например, в 2006 г. американский журнал *Newsweek* подготовил свой глобальный рейтинг университетов «The Top 100 Global Universities», исходной информационной базой для которого служили два авторитетных рейтинга – ARWU и THE. Тем самым на рынке глобальных рейтингов у двух «материнских» рейтинговых систем как бы появился «дочерний» рейтинговый продукт, что автоматически усилило их позиции. Нечто похожее произошло в 2008 г., когда под эгидой U.S. News & World Report's на рынок вышел американский рейтинг World's Best Universities ranking (WBU). Начиная с 2010 г. этот рейтинг базируется на данных рейтинга QS и активно сотрудничает с его разработчиком – компанией «Quacquarelli Symonds». Иными словами, рейтинг QS осуществил ответные действия и вывел на международный рынок свой «дочерний» продукт.

В этих примерах мы видим, что британские фирмы входят в альянс с американскими компаниями, тем самым распространяя влияние своей методологии и своих данных на разработчиков из другого региона. Таким образом, в табл. 2 рейтинговая система QS присутствует неявно – через рейтинг WBU. «Вклинивание» в этот процесс китайского оператора ARWU еще больше усложняет истинный «расклад сил» между рейтингами.

Альянсы между рейтинговыми агентствами дополняются и политикой диверсификации рейтинговых продуктов, о чем упоминалось выше. Например, в 2011 г. новый рейтинг World Reputation Rankings (WRR), разработанный «Times Higher Education» с агентством *Thomson Reuters*, позволил этим компаниям стать лидерами на совершенно новом сегменте рынка – в области «чистых» персонализированных глобальных рейтингов. Такой шаг в очередной раз дает конкурентные преимущества британскому бренду THE.

4. Оценка взаимной согласованности глобальных рейтингов

Конкурентная борьба рейтингов, ведущаяся рыночными методами, имеет и чисто научное измерение. Например, можно ли определить, какой из существующих рейтингов более объективный и точный?

Однако прежде чем пытаться проверить глобальные рейтинги в смысле их достоверности, выясним то, насколько они согласованы между собой. Данная работа уже проводилась, в частности, в исследовании (Aguillo et al., 2010), в котором была сделана попытка определить *согласованность* разных глобальных рейтингов по группам университетов – Топ-10, Топ-100 и Топ-200 (Aguillo, Bar-Ilan, Levene, Ortega, 2010). Ниже мы попытаемся проделать эту работу, но для иного массива глобальных рейтингов. Так, в указанной работе использовалось 5 рейтингов (ARWU, THE-QS, Web, PRSP и LR) – мы же рассмотрим 6 рейтингов (табл. 3).

В основной серии расчетов нами использовался массив данных Топ-25 университетов. Ранги (места) университетов оценивались попарно между всеми рейтингами с помощью обычного коэффициента корреляции³. Результаты расчетов приведены в табл. 3, в которой полужирным шрифтом обозначены значимые коэффициенты корреляции с коэффициентом значимости в 1%.

Результаты расчетов позволяют сделать следующие выводы.

Единственным самостоятельным и не связанным с другими рейтингами является тайваньский PRSP. Почти автономным является Web, он слабо согласуется только с ARWU. Аналогично выглядит рейтинг QS, имеющий явную согласованность с THE, с которым он связан исторически и методологически (ранее они составляли единый рейтинг QS-THES), и с WRR, который также разрабатывается их бывшей общей «материнской» компанией. Именно эти три рейтинга – PRSP, Web и QS – можно считать маргинальными в том смысле, что их результаты плохо коррелируют с другими глобальными рейтингами.

Таблица 3

Взаимная корреляция глобальных рейтингов университетов

Рейтинги	QS	PRSP	ARWU	THE	Web	WRR
QS	1	-0,030	0,317	0,601	-0,215	0,613
PRSP	-0,030	1	0,300	0,150	0,195	0,225
ARWU	0,317	0,300	1	0,607	0,493	0,529
THE	0,601	0,150	0,607	1	0,049	0,652
Web	-0,215	0,195	0,493	0,049	1	-0,041
WRR	0,613	0,225	0,529	0,652	-0,041	1

³ В наших расчетах использовался обычный коэффициент корреляции, хотя по логике исследования следовало бы воспользоваться каким-то коэффициентом ранговой корреляции (например, коэффициентом Спирмена). Однако это невозможно в силу специфики используемых массивов данных. Дело в том, что если в одном массиве (рейтинге) из 25 университетов мы имеем строго мажорированное упорядочение ранга от 1 до 25, то в других массивах (рейтингах) нарушается не только процедура мажорирования, но могут возникать ранги с числом больше 25, причем некоторые внутренние значения ряда [1; 25] могут выпадать. В этом случае использование рангового коэффициента корреляции становится некорректным.

Остальные три рейтинга – THE, WRR и ARWU – имеют высокий статус согласованности, равный трем⁴. При этом самыми высокими коэффициентами корреляции отличается рейтинг THE, за ним идет WRR и замыкает группу ARWU.

Учитывая, что мы сравниваем два традиционных, два гибридных, один виртуальный и один репутационный рейтинги, наши результаты представляются вполне объективными, так как в них не прослеживается явного преимущества двух пар. Наоборот, в каждой паре имеются лидеры (THE и ARWU) и аутсайдеры (QS и PRSP). Таким образом, вхождение в определенный класс рейтинга еще не гарантирует сходства рейтингов. Особенно ярко это видно на примере ARWU, который фактически не согласуется с «родственным» рейтингом PRSP, но довольно хорошо согласуется с рейтингами других классов – THE, WRR и Web (см. табл. 3).

Любопытно, что наши расчеты не выявили устойчивой группы взаимосогласованных рейтингов. Так, три рейтинга, претендующих на роль лучших (THE, WRR и ARWU), не образуют замкнутой группы; в нее «примешиваются» Web и QS. Таким образом, явная *групповая согласованность глобальных рейтингов отсутствует*, что делает проблему оценки их надежности еще более насущной и нетривиальной.

Во второй серии расчетов нами использовался массив данных для пяти рейтингов (без WRR) для университетских выборок Топ-25, Топ-50 и Топ-75. Укажем основные результаты. Так, для Топ-25 результаты оказались практически идентичными данным табл. 3. Однако увеличение выборки приводит к росту согласованности рейтингов. Например, для Топ-50 рейтинг PRSP оказался согласованным уже с двумя рейтингами – QS и ARWU. В Топ-75 практически все рейтинги оказались согласованными между собой; не просматривалось согласованности только между QS и Web. Таким образом, рост выборки увеличивает согласованность глобальных рейтингов. По-видимому, это связано с тем, что для больших выборок более значимым становится не столько характер ранжирования, сколько сам факт ранжирования элементов. Иными словами, на больших выборках качественная похожесть рейтингов лучше заметна, чем на малых.

5. Методика анализа надежности глобальных рейтингов

Все существующие на сегодняшний день рейтинги чрезвычайно уязвимы. Однако даже подробный анализ методов манипулирования с ними не позволяет прийти к выводу о том, какой же из них является наиболее предпочтительным. Между тем их множественность сама по себе открывает новые возможности для анализа степени их адекватности. Опираясь на этот вывод, рассмотрим простую схему, которая позволяет внести некоторую дополнительную ясность в вопрос об их достоверности.

Наш алгоритм представляет собой эвристическую процедуру, состоящую из нескольких простых шагов.

⁴ Под статусом согласованности данного рейтинга нами понимается число других рейтингов, с которыми у него имеется значимый коэффициент корреляции. Так, у PRSP данный статус равен 0, у Web – 1, у QS – 2, у THE, WRR и ARWU – 3.

На *первом шаге* вычисляется медианный, т.е. усредненный, глобальный рейтинг университетов:

$$z_j = \frac{1}{m} \sum_{i=1}^m x_{ij}, \quad (1)$$

где i – индекс рейтинга; j – индекс университета; $j = 1, \dots, N$; N – число анализируемых университетов; m – число анализируемых рейтингов; x_{ij} – ранг (место) университета j в рейтинге i ; z_j – ранг (место) университета j в медианном рейтинге.

В основе данной процедуры лежит простая идея о том, что при объединении различных рейтингов «включается» закон больших чисел⁵. Тем самым усредненная оценка по многим испытаниям (рейтингам) дает некий ориентир истинного места изучаемого университета. Подчеркнем, что полученный медианный рейтинг имеет вспомогательное значение, поэтому его не следует воспринимать в качестве самостоятельного индикатора. Вектор (1) выступает в качестве *квази-объективной оценки* университетских рангов.

На *втором шаге* рассчитывается показатель среднего отклонения δ_i рангов конкретного рейтинга от оценок медианного рейтинга (1):

$$\delta_i = \frac{1}{N} \sum_{j=1}^N |x_{ij} - z_j| \quad (2)$$

На основе показателя (2) можно сравнить степени адекватности глобальных рейтингов. Если δ_i имеет незначительную величину, то это означает, что ранги рейтинга i не намного отклоняются от рангов медианного рейтинга и, следовательно, результаты этих двух рейтингов хорошо согласованы. В этом случае подобный рейтинг вправе претендовать на высокий уровень доверия. В противном случае, когда ранги рейтинга i существенно отклоняются от рангов медианного рейтинга, можно говорить об их рассогласовании, и рейтинг i следует классифицировать как ненадежный.

На *третьем шаге* предстоит определить, какие рейтинги лежат в зоне допустимых отклонений, а какие выходят за ее пределы. Будем считать, что рейтинг обладает свойством *глобальной надежности*, если выполняется условие:

$$\delta_i < \delta_{\text{крит}}, \quad (3)$$

где $\delta_{\text{крит}}$ – критическое (допустимое) значение δ_i .

При нарушении правила (3) рейтинг будем считать глобально ненадежным. Для определенности будем полагать, что

$$\delta_{\text{крит}} = N/3. \quad (4)$$

В основе оценки (4) лежит следующая логика. Любую изучаемую группу университетов можно механически разбить на три равные подгруппы. Первая подгруппа будет состоять из «хороших» (лидеров) вузов, вторая подгруппа – из «средних», а третья – из «плохих» (аутсайдеров). Тогда правомерно предположить, что недопустимым должно

⁵ Не следует, однако, буквально воспринимать применение закона больших чисел к глобальным рейтингам. Строго говоря, наличие 5–6 рейтингов еще не «включает» закона больших чисел, но, учитывая значимость и конструкцию рейтингов, такой подход представляется правомерным.

считаться такое отклонение, при котором университет может переходить в качественно другую подгруппу, т.е. осуществлять переход через подгруппу. Например, вполне допустимо, что университет из лидеров перейдет в подгруппу середняков, но будет странно, если он окажется в подгруппе отстающих. Данная логика справедлива и в обратном направлении, когда аутсайдер вдруг окажется в группе лидеров. Подобные качественные рассогласования будем считать недопустимыми. Учитывая, что (2) показывает среднее число мест (рангов), на которое отклоняется рейтинг вуза от своего медианного значения, его значение не должно превышать длину подгруппы – $N/3$; в противном случае будет возможно перескакивание через подгруппу середняков. Этим обстоятельством и предопределяется значение (4).

Помимо выводов относительно глобальных рейтингов можно сделать аналогичные выводы и для университетов, входящих в анализируемую группу. Например, в целом рейтинг может обладать свойством глобальной надежности, но при этом ряд вузов он оценивает все-таки не очень удачно. В этом случае можно говорить о *локальной надежности* в том смысле, что некоторые вузы из рассматриваемой выборки обладают избыточной подвижностью, не вписываясь в рамки установленных ограничений. Тогда можно применить процедуру просеивания вузов по принципу: если для университета j в рейтинге i выполняется условие $|x_{ij} - z_j| < N/3$ (в дальнейшем данное неравенство будем называть условием локальной надежности), то этот университет обладает свойством локальной надежности; в противном случае вуз не обладает указанным свойством.

Результат просеивания университетов может быть зафиксирован с помощью следующего долевого показателя:

$$\mu = N^* / N, \quad (5)$$

где N^* – число университетов в выборке, которые удовлетворяют условию локальной надежности.

Процедура просеивания университетов позволяет, во-первых, перейти к микроанализу при изучении конкретных вузов, в отношении которых имеются вопросы о правомерности их ранга, а во-вторых, получить дополнительный показатель надежности изучаемого рейтинга. Для более тонкого анализа показателя μ можно использовать условие, согласно которому допустимым считается доля локально ненадежных университетов в $1/3$ изучаемой выборки. Тогда рейтинг считается локально надежным при $\mu > 66,6\%$.

В дальнейшем показатель (2) будем считать коэффициентом глобальной надежности рейтинга, а коэффициент (5) – коэффициентом локальной надежности.

6. Информационная база расчетов

Предложенный подход к оценке надежности рейтингов университетов был апробирован на данных рассмотренных ранее наиболее

репрезентативных рейтингов. Нами были проведены две серии расчетов. В первой использовалось 6 рейтингов ($m = 6$), во второй – только 5 ($m = 5$). При этом все рейтинги использовались за 2010 г., а рейтинг WRR – за 2011 г. Добавление к пяти синхронизированным рейтингам дополнительного рейтинга за другой год объясняется тем обстоятельством, что рейтинг WRR является достаточно консервативным и быстро не меняется, а также тем, что он был составлен на основе опросов 2010 г. и тем самым отражал ситуацию в основном за предыдущий год. Следовательно, содержание шестого рейтинга полностью соответствует остальным рейтингам. Эта логика распространяется и на предыдущие расчеты взаимной согласованности рейтингов.

В первой серии расчетов использовалась только одна короткая выборка из 25 университетов, во второй серии – три выборки из 25, 50 и 75 вузов. Такие ограничения в выборках были связаны со сложностью «стыковки» вузов разных рейтингов. Например, вуз, стоящий в середине первой сотни в одном рейтинге, может оказаться в середине второй сотни в другом рейтинге. Таким образом, речь идет не о простой перестановке университетов внутри выборки, а об их «рокировке» в рамках некоторой расширенной выборки. Соответственно, ранги многих университетов далеко выходят за пределы изучаемой выборки.

Использование только одной 25-элементной выборки для первой серии расчетов связано с тем, что рейтинг WRR после 25 места очень плохо соответствует другим рейтингам, будучи самым коротким рейтингом из всех существующих (только Top-100); его ограниченная информационная база не позволяла найти университеты, необходимые для сравнения по другим рейтингам.

Вторая серия расчетов проводилась для трех выборок, однако уже даже для выборки в $N = 100$ оказалось невозможно найти все нужные вузы в открытых базах изучаемых рейтингов.

Плохая стыковка университетов в разных выборках рейтингов привела к тому, что проведение сравнительных расчетов потребовало определения некоего реперного (опорного) рейтинга, который задает исходный массив университетов для изучаемой выборки. Такой массив выбирается в значительной степени произвольно и субъективно. В качестве реперного рейтинга нами был выбран ARWU – как самый «старый» и один из самых авторитетных рейтингов. На первый взгляд, может показаться, что такой особый статус ARWU дает ему преимущества в последующих вычислениях. Однако на самом деле, даже если это и происходит, то «вычислительный бонус» ARWU, будучи случайной величиной, пренебрежимо мал, и им можно реально пренебречь без потери «содержания» результатов.

7. Результаты эмпирических расчетов

Результаты расчетов приведены в табл. 4–5, анализ которых позволяет сделать ряд предварительных и очень осторожных выводов.

Так, в первой серии расчетов ($m = 6$) хорошо просматривается иерархия рейтингов по уровню надежности (табл. 4). При этом три из них – WRR, THE и ARWU – проходят тесты на глобальную и локальную надежность. Остальные три рейтинга – PRSP, QS и Web – по обоим тестам должны быть признаны недостаточно надежными. Тем самым лучшими из существующих глобальных рейтингов университетов оказываются два самых старых из них (THE и ARWU) и репутационный рейтинг WRR, который занял лидирующее положение. Судя по всему, именно этим рейтингам и следует доверять больше всего; именно на них и следует ориентироваться.

Заметим, что этот результат вполне соответствует данным о популярности рейтингов (см. табл. 2). Тем самым мы имеем довольно редко встречающуюся картину: наиболее надежные рейтинги одновременно являются и наиболее популярными.

Наш вывод, однако, нуждается в некоторой корректировке. Дело в том, что репутационный рейтинг WRR может считаться «хорошим» только для очень короткого списка рейтинга, когда оценке подвергаются только 25 университетов-лидеров. Для более широкого списка этот рейтинг становится сомнительным, если не сказать, бессмысленным. Это свойство WRR непосредственно связано с недостатками в методике его построения. Иными словами, мы определили пределы, в которых репутационные рейтинги являются высокоэффективными⁶.

Результаты второй серии расчетов подтверждают наши выводы (табл. 5). Два «старых» рейтинга – THE и ARWU – по-прежнему сохранили лидерство. Более того, британский рейтинг и здесь опережает шанхайский. Несколько нарушает картину только тайваньский рейтинг PRSP, который в выборке с $N = 75$ по обоим критериям попал в группу надежных рейтинговых систем. Следовательно, можно говорить о том, что на большом массиве статистики данный рейтинг претендует на статус вполне надежного измерителя. Остальные два рейтинга – QS и Web – неизменно сохраняют позиции маргинальных систем.

Следует отметить, что рейтинг WRR хорошо работает в начале рейтинга и очень плохо на его «теле», тогда как PRSP – наоборот, он дает недостоверные оценки на довольно длинном «хвосте» вначале и вполне приличные – на остальной части. Тем самым можно говорить о «короткой» и «длинной» специализации глобальных рейтингов.

Важной закономерностью второй серии расчетов является рост отклонения (2) по мере

Таблица 4

Характеристики глобальных рейтингов ($m = 6$)

Рейтинги	$N = 25; \delta_{\text{крит}} = 8,3$	
	δ	μ
THE	6,8	76,0
ARWU	6,6	68,0
WRR	7,1	76,0
PRSP	10,1	56,0
QS	12,9	52,0
Web	21,5	36,0

⁶ Подчеркнем, что зафиксированный предел эффективности в 25 вузов является совершенно нетривиальным и неожиданным. На основе распределения рейтинговых оценок было вполне логично предположить, что этот предел гораздо уже и не превышает 10–12 позиций.

Таблица 5
Характеристики глобальных рейтингов ($m = 5$)

Рейтинги	$N = 25; \delta_{\text{крит}} = 8,3$		$N = 50; \delta_{\text{крит}} = 16,7$		$N = 75; \delta_{\text{крит}} = 25$	
	δ	μ	δ	μ	δ	μ
THE	7,5	68,0	15,0	76,0	18,5	74,7
ARWU	7,8	68,0	16,1	70,0	22,9	66,7
PRSP	10,4	52,0	18,4	60,0	24,5	70,7
QS	13,7	40,0	27,0	46,0	37,4	46,7
Web	20,5	36,0	31,7	52,0	50,7	48,0

роста выборки: $\partial\delta_i / \partial N > 0$. Тем самым при увеличении числа сравниваемых университетов все рейтинги как бы «расплываются», становясь все менее точными. Если даже в выборке в 100 вузов величина среднего отклонения для THE изменится незначительно, то можно предполагать, что она составит около 20 позиций. Это означает, например, что университет с рангом 50 может занимать от 30 до 70-го места. А это – очень низкая «точность». Таким образом, выявленная закономерность ($\partial\delta_i / \partial N > 0$) свидетельствует о принципиально ограниченной индикативной роли глобальных рейтингов университетов. Фактически это означает, что все глобальные рейтинги по сути позволяют ранжировать только университеты-лидеры. За пределами первой сотни потенциальная ошибка возрастает настолько, что сами рейтинги становятся условными, поэтому их можно использовать лишь для выявления качественных моментов в развитии вуза. Например, университет вошел в первую, вторую или третью сотню вузов.

Интересно, что в отношении коэффициента μ столь простой закономерности не просматривается. Здесь имеет место следующая картина: для «хороших» рейтингов локальная эффективность по мере роста выборки сначала увеличивается, а потом снижается; для «плохих» рейтингов прослеживается однонаправленная тенденция роста μ . Таким образом, можно лишь утверждать, что появление университетов-маргиналов в разных выборках различных рейтингов происходит неодинаково.

Было бы неверно абсолютизировать результаты расчетов. Однако они все-таки вносят некоторую ясность в диспозицию существующих глобальных рейтингов университетов, что само по себе имеет немаловажное значение.

Литература

- Арский Ю.М., Маркусова В.А., Чумакова Н.Ф.** (2007): Науки о Земле: объемы и рейтинги отечественных публикаций // *Вестник Российской академии наук*. Т. 77. № 10.
- Балацкий Е.В., Екимова Н.А.** (2011): Международные рейтинги университетов: практика составления и использования // *Журнал Новой экономической*

ассоциации. № 9.

- Гинзбург В.Л.** (1998): Почему советские ученые не всегда получали заслуженные ими Нобелевские премии? // *Вестник Российской академии наук*. Т. 68. № 1.
- Кларк М.** (2002): Некоторые идеи о рейтингах академического качества // *Высшее образование в Европе*. Т. XXVII. № 4. Режим доступа: <http://technical.bmstu.ru/istoch/kach/idorejt.doc>, свободный. Загл. с экрана. Яз. англ. (дата обращения: декабрь 2010 г.).
- Салми Дж.** (2009): Создание университетов мирового класса. М.: Весь мир.
- Свердлов Е.Д.** (2006): Миражи цитируемости. Библиометрическая оценка значимости научных публикаций отдельных исследователей // *Вестник Российской академии наук*. Т. 76. № 12.
- Aguillo I.F., Bar-Ilan J., Levene M., Ortega J.L.** (2010): Comparing University Rankings // *Scientometrics*. Vol. 85. February.
- Arnold D.N., Fowler K.K.** (2011): Nefarious Numbers // *Notices of the AMS*. Vol. 58. № 3.
- Baty Ph.** (2010): Measured, and Found Wanting More [Электронный ресурс] // *Times Higher Education*. 8 July. Режим доступа: <http://www.timeshighereducation.co.uk/story.asp?storycode=412341>, свободный. Загл. с экрана. Яз. англ. (дата обращения: апрель 2011 г.).
- Diamond N., Graham H.** (2000): How Should We Rate Research Universities? [Электронный ресурс] Режим доступа: <http://www.pha.jhu.edu/~sujian/ranking.htm>, свободный. Загл. с экрана. Яз. англ. (дата обращения: апрель 2011 г.).
- Holmes R.** (2010): The THE-QS World University Rankings, 2004–2009. Universities Rankings Watch, October, 19. Режим доступа: <http://www.universities06.com/theqsranking.pdf>, свободный. Загл. с экрана. Яз. англ. (дата обращения: апрель 2011 г.).
- Marginson S.** (2007): Global University Rankings: Where to from Here? / Asia-Pacific Association for International Education, National University of Singapore, 7–9 March. Режим доступа: http://www.cshe.unimelb.edu.au/people/marginson_docs/APAIE_090307_Marginson.pdf, свободный. Загл. с экрана. Яз. англ. (дата обращения: май 2010 г.).
- Smith D.N.** (2010): The History Behind Universities, League Tables and the Brand [Электронный ресурс] // *Global Opinion Survey. New Outlooks on Institutional Profiles*. February. Режим доступа: http://science.thomsonreuters.com/m/pdfs/Global_Opinion_Survey.pdf, свободный. Загл. с экрана. Яз. англ. (дата обращения: декабрь 2010 г.).
- The Leiden Ranking 2008 (2008):** [Электронный ресурс] Рейтинг. Голландия. Режим доступа: <http://www.cwts.nl/ranking/LeidenRankingWebSite.html>, свободный. Загл. с экрана. Яз. англ. (дата обращения: май 2011 г.).
- The Times Higher Education (2010a):** [Электронный ресурс] Рейтинг. Великобритания. Режим доступа: <http://www.timeshighereducation.co.uk/world-university-rankings/2010-2011/reputation-methodology.html>, свободный. Загл. с экрана. Яз. англ. (дата обращения: апрель 2011 г.).

The Times Higher Education (2010b): [Электронный ресурс] Рейтинг. Великобритания. Режим доступа: <http://www.timeshighereducation.co.uk/world-university-rankings>, свободный. Загл. с экрана. Яз. англ. (дата обращения: апрель 2011 г.).

Van Raan A.F.J. (2005): Challenges in Ranking Universities/ Invited paper for the First International Conference on World Class Universities, Shanghai Jiao Tong University, Shanghai. June 16–18.

Usher A., Savino M. (2007): A Global Survey of Rankings and League Tables. College and University Ranking Systems. Global Perspectives and American Challenges/ Institute for Higher Education Policy (editor). Режим доступа: <http://www.educationalpolicy.org/pdf/World-of-Difference-200602162.pdf>, свободный. Загл. с экрана. Яз. англ. (дата обращения: апрель 2011 г.).

U.S. News & World Report (2011): [Электронный ресурс] Рейтинг. США. Режим доступа: <http://www.usnews.com>, свободный. Загл. с экрана. Яз. англ. (дата обращения: апрель 2011 г.).

Поступила в редакцию 2 мая 2011 г.

E.V. Balatsky
CEMI RAS, Moscow

N.A. Ekimova
The State University of Management, Moscow

The Comparative Reliability of the Global University Rankings

The plurality of global rankings of Universities is considered in this article. The classification of global ranking is given. Their mutual consistency is evaluated. The procedure of estimation and comparison of the reliability of global rankings is offered and approved.

Keywords: *rankings of Universities, ranking criteria, education, economic policy, local reliability, global reliability.*

JEL classification: I21, I23.