

МОДЕЛИРОВАНИЕ УСТОЙЧИВОСТИ РОССИЙСКИХ БАНКОВ В ПЕРИОД РЕФОРМИРОВАНИЯ БАНКОВСКОЙ СИСТЕМЫ

К.Л. Поляков,
М.В. Полякова,

Авторами статьи излагаются результаты исследования, актуального в условиях реформирования отечественной банковской системы, по математико-статистическому моделированию устойчивости функционирования банков. Концептуальная основа таких построений - понимание устойчивости (или надежности) функционирования предприятия, финансовое состояние которого обеспечивает в нормальных условиях выполнение всех его обязательств перед работниками, другими организациями, государством благодаря достаточным доходам, соответствию доходов и расходов. Некоторые из значимых для анализа устойчивости (надежности) банков индикаторов доступны внешним заинтересованным лицам и организациям за счет раскрытия большинством банков своей ежемесячной финансовой отчетности, которая размещается и регулярно обновляется на сайте Центрального банка Российской Федерации. К числу таких показателей, по мнению авторов, относятся показатели, для которых Банк России определяет нормативные значения. Хотя отслеживание степени соответствия фактических индикаторов нормативам представляется далеко не идеальным инструментом анализа надежности банка (для диагностирования ситуации, не приводящей к отзыву лицензии), значения данных показателей, по мнению авторов, являются достаточными характеристиками финансового здоровья кредитного учреждения.

В публикуемой работе исследовался характер статистической взаимосвязи между трендами количественных значений параметров, характеризующих важнейшие стороны деятельности банков, с вероятностью отзыва у них лицензии. Спецификой построения моделей является автоматический выбор функциональной формы вхождения в них характеристик банков. С этой целью авторы используют аппарат обобщенных полиномов, которые позволяют выбрать спецификацию модели, наиболее адекватную свойствам данных.

Исследование проведено на основании открытой отчетности 887 банков за период с 01.01.2013 по 01.12.2015. Для анализа использовалась отчетность как действующих, в частности начавших работу, так и ликвидированных в этот период банков (некоторые из действующих в этот период банков позже - в 2016 и 2017 гг. также лишились лицензии). Результаты оценивания моделей показали высокую степень согласования формы вхождения в них характеристик банка с их экономическим смыслом и нормативами Центрального банка Российской Федерации. Сравнение качества классической модели бинарного выбора для панельных данных с моделью, основанной на обобщенных полиномах, показывает явное преимущество последней.

Ключевые слова: устойчивость банка, отзыв лицензии, отчетность банков, банковская статистика, модель бинарного выбора, панельные данные, обобщенные полиномы.

JEL: C33, C51, G01, G18, G21, G22, G28.

Трехлетний период с 2013 по 2015 г. можно охарактеризовать как период реформирования российской банковской системы. Усиление строгости контроля Центрального банка Российской Федерации за деятельностью участников рынка, в частности за соблюдением нормативов величины капитала, ликвидности и т. д., сопровождалось введением ряда существенных инноваций, повлиявших на «правила игры» в сегменте финансовых услуг. К ним можно отнести введение ключевой ставки в сентябре 2013 г. и интенсивное управление ее величиной, а также двукратное увеличение страхового возмещения в Системе страхования вкладов и ее распространение на счета индивидуальных предпринимателей 19 декабря 2014 г.

Данное исследование посвящено анализу устойчивости российских банков в указанный выше период, где под устойчивостью понимается: «финансовое состояние предприятия, хозяй-

ственная деятельность которого обеспечивает в нормальных условиях выполнение всех его обязательств перед работниками, другими организациями, государством благодаря достаточным доходам и соответствию доходов и расходов» [1]. В литературе, связанной с банковским бизнесом, чаще используется аналогичное понятие «надежность».

На устойчивость (надежность) банка влияет множество факторов, в частности организационных, многие из которых могут быть проанализированы только в рамках процедуры аудита организации [2]. Однако некоторые из значимых для анализа устойчивости (надежности) банков факторов доступны внешним заинтересованным лицам и организациям за счет раскрытия большинством банков своей ежемесячной финансовой отчетности, которая размещается и регулярно обновляется на сайте Центрального банка Рос-

Поляков Константин Львович (polyakov.kl@hse.ru) - канд. техн. наук, доцент, департамент прикладной экономики факультета экономических наук, Национальный исследовательский университет «Высшая школа экономики» (г. Москва, Россия).

Полякова Марина Васильевна (mpolyakova@hse.ru) - канд. техн. наук, доцент, Школа финансов факультета экономических наук, Национальный исследовательский университет «Высшая школа экономики» (г. Москва, Россия).

сийской Федерации. К числу таких показателей относятся обязательные нормативы банков РФ¹. Несмотря на то, что, как правило, с ними связывают контроль за деятельностью кредитных организаций со стороны Банка России, значения этих нормативов с точки зрения авторов являются полноценными характеристиками финансового здоровья.

В данном исследовании ставится задача анализа характера статистической взаимосвязи значений указанных выше показателей с вероятностью отзыва лицензии. При этом в центре внимания находятся сила и направление влияния, а также возможная его смена. Это дает возможность менеджменту банка повысить оперативность и качество принятия решений о его управлении, а партнерам и клиентам - выполнять внешний экспресс-анализ надежности банка, используя относительно небольшой набор показателей. Отметим, что авторы концентрируют внимание на описании функциональной формы статистической взаимосвязи и не ставят задачу построения в некотором смысле наилучшей системы прогнозирования отзыва лицензии. Актуальность и различие этих двух задач рассмотрены во многих публикациях, в частности в [3, 4]. Для решения поставленной задачи в работе используется сочетание моделей бинарного выбора для панельных данных, функция индекса которых формируется с помощью обобщенных полиномов (Fractional Polynomials).

Структура работы следующая. В разделе «Показатели и модели» представлен обзор публикаций, в которых можно найти важные для настоящего исследования результаты, связанные как со спецификой российского рынка банковских услуг, так и с классом используемых моделей. В разделе «Описание используемых данных» приводится обзор множества банков, включенных в анализ, и показателей, характеризующих их устойчивость. Раздел «Математические основы исследования» посвящен краткому описанию моделей, которые использовались для решения поставленной в исследовании задачи.

Показатели и модели

Теме анализа устойчивости банков посвящено немало работ. Общие подходы к решению этой задачи сформулированы, в частности, в работе [5] в контексте оценки рисков и формирования

«систем раннего предупреждения» в странах Большой десятки (The Group of Ten, G10). Анализируя существующую практику организации надзора и регулирования в банковской деятельности, выделяют четыре основные категории подходов, одной из которых является статистическое моделирование «здоровья» банков, которому посвящено настоящее исследование. Теоретическое обоснование подходов из этой категории получило развитие в работе [6]. Авторы отчета формализуют понятие «слабый банк» (weak bank):

«Слабым банком называется банк, чья ликвидность или платежеспособность могут серьезно пострадать, если не будут значительно улучшены его финансовые ресурсы, профиль рисков, стратегия развития, возможности риск менеджмента и/или качество управления».

При этом важно уметь различать слабые банки и банки, где проблемы имеют локальный характер и могут быть относительно быстро ликвидированы адекватными мерами. Слабые банки не возникают в режиме «overnight», негативные тенденции развиваются на достаточно протяженных отрезках времени как результат слабого управления. Важнейшей задачей регулятора является обнаружение и своевременная корректировка подобных ситуаций. Подходы к ее решению существенно зависят от национального законодательства, но в любом случае от регулятора требуется поддержка стабильности и надежности финансовой отрасли за счет снижения риска потерь для любых кредиторов. Отметим, что именно эта задача решается ЦБ РФ в период «очистки финансового сектора», которому посвящено настоящее исследование.

Вместе с тем анализ «слабости» банка актуален не только для регулятора. Участники многих сегментов рынка заинтересованы в наличии надежного финансового партнера для реализации средне- и долгосрочных проектов. При этом в большинстве случаев они не обладают возможностями государственных надзорных органов или крупных аналитических компаний в отношении доступа к закрытым внутренним данным кредитных организаций и должны довольствоваться открытыми данными и ориентированными на них методиками оценки уровня риска при взаимодействии с банком. Как правило, при решении данной задачи используются различные методики оценки кредитного риска [7-9]. Одной из наиболее распространенных является задача

¹ Инструкция об обязательных нормативах банков // Вестник Банка России. 2012. № 74. URL: <http://www.cbr.ru/publ/Vestnik/>.

оценки вероятности наступления дефолта, то есть неспособности заемщика выполнить свои обязательства в полном объеме в срок. Однако при решении практических задач определение этого понятия варьируется, что следует учитывать при сравнении результатов в различных исследованиях. Например, популярная модель Альтмана разработана для прогнозирования банкротства, а не менее популярная модель Чессера позволяет анализировать вероятность отклонения от условий кредитного договора [8]; при этом обе ситуации трактуются как дефолт.

В российской практике банковской аналитики для анализа надежности банков широкое распространение получили модели вероятности отзыва лицензии, в первую очередь в связи с наступлением периода интенсивных реформ сегмента финансовых услуг. Сам по себе факт отзыва лицензии формально не влечет за собой потерь для клиентов банка - как кредиторы они в соответствии с законодательством могут рассчитывать на первоочередной возврат своих средств. Однако если величина активов банка на момент закрытия меньше его обязательств, потери неизбежны. Поэтому то или иное решение данной задачи позволяет получить оценку величины рисков при работе с данной кредитной организацией, ее надежности. Большинство исследований в этой области основано на использовании моделей бинарного выбора и могут быть охарактеризованы конкретными версиями этих моделей, составом независимых переменных, а также конкретизацией понятия «дефолт». Впервые модели этого типа были использованы для решения рассматриваемой задачи в работе [10].

Одно из наиболее ранних исследований надежности российских банков представлено в работе [11]. Авторы на основе опроса экспертов отобрали несколько показателей, в наибольшей степени характеризующих надежность банка, и несколько наиболее надежных рейтинговых агентств. Далее на основе реальных данных банковской отчетности были построены модели статистической взаимосвязи выбранных показателей с оценками отобранных агентств. Использовались модели множественного выбора. Отметим, что множества банков для разных агентств несколько различались. Использовались данные за осень 2001 г. Интересно, что во всех построенных моделях наиболее значимыми показателями, направление влияния которых согласуется со здравым смыслом, оказались показатели, для значений

которых Банк России устанавливает нормативы: показатель достаточности капитала (норматив Н1), показатели ликвидности (нормативы Н2, Н3, Н5). Ни одна из моделей не показала значимости доли просроченной задолженности в кредитах.

Также важной для настоящего исследования является работа [12]. Авторы вводят понятие банкротства как факта отзыва лицензии, не связанного с реструктуризацией, или перехода банка под управление Агентства по реструктуризации кредитных организаций (АРКО). В исследовании ставится задача определения показателей, статистически значимо связанных с фактом банкротства. Для этого используются логит-модели бинарного выбора. Авторы предполагают, что множество банков в исследовании неоднородно в контексте данной задачи, то есть сила и направление влияния характеристик банка на вероятность банкротства могут считаться постоянными только в границах отдельных кластеров. Фактически в исследовании предполагается возможная нелинейная связь выбранных характеристик и вероятности банкротства. Для решения поставленной задачи авторы предлагают две авторские альтернативные процедуры кластеризации. Для полученных кластеров были построены модели указанного выше типа; в ряде случаев некоторые независимые переменные включались в модели в первой и второй степенях. Сравнение результатов моделирования показало, что предположение о неоднородности множества банков в контексте поставленной задачи вполне обоснованно. Похожие гипотезы проверяются в работе [13]. Авторы этого исследования анализировали динамику структуры российского банковского сектора, где под структурой понимается кластерная модель множества российских банков, основанная на данных ежегодных открытых отчетов банков, размещенных на сайте Центрального банка Российской Федерации. В исследовании с использованием анализа трендов выделяется несколько периодов развития российского рынка банковских услуг. В результате авторы выделяют несколько статистически значимо различающихся типов банков с точки зрения устойчивости, которые встречаются на каждом из указанных этапов. Эти работы свидетельствуют о необходимости учета нелинейного влияния различных индикаторов деятельности банка на его устойчивость, в частности на вероятность отзыва лицензии.

В работе [14] детализируется понятие дефолта. На основе анализа приказов Центрального банка

Российской Федерации автор выделяет несколько категорий причин отзыва лицензии и для каждой из них определяет наиболее существенные для оценки вероятности отзыва лицензии факторы. Для решения данной задачи используются модели дискретного выбора (бинарного и множественного), выделяются факторы, имеющие значимые оценки коэффициентов в функции индекса модели. Как показало исследование, группы существенных факторов для разных категорий причин отзыва лицензии значительно различаются. Отметим, что определяя категории причин отзыва лицензии, автор опирается прежде всего на наличие/отсутствие чисто экономических причин и нарушений Федерального закона № 115. С нашей точки зрения, это полезное предложение, поскольку, как показывает анализ отчетности банков, отзыв лицензии по «криминальным причинам» часто осуществляется в отношении банков, которые с экономической точки зрения весьма благополучны, в частности соблюдаются все нормативы Центрального банка Российской Федерации. Таким образом, очевидно, что оценка вероятности дефолта должна осуществляться на основе разных, но, возможно, пересекающихся групп показателей. В начальное множество факторов были, помимо прочего, включены показатели кредитной активности, характеристики ликвидности активов, отношение собственного капитала к чистым активам. Характерно, что все показатели рассматривались как отношения к чистым активам. По нашему мнению, использование этой характеристики более корректно с экономической точки зрения, чем валюты баланса (суммарных активов), поскольку часть сумм, проходящих по балансу, как активы, на самом деле не способны приносить прибыль и не могут использоваться для оценки эффективности и устойчивости банка.

Работа [15] развивает использование моделей конечного выбора для оценки вероятности дефолта банка. Авторы вводят очень подробное определение понятия «дефолт», используя анализ экономических причин отзыва лицензии. Необходимо отметить, что оно существенно отличается от определений в разобранных выше работах. Для оценки вероятности дефолта в данном исследовании используется стандартная логит-модель. Спецификой работы является тщательная очистка базы данных от наблюдений, которые оценены как сомнительные. Кроме того, был проведен анализ коррелированности коэффициентов, и

из их числа были исключены некоторые, сильно коррелирующие с остальными. Так же, как и в предыдущей работе, для формирования коэффициентов, характеризующих состояние банка, используются чистые, а не совокупные активы. Отметим, что в данном исследовании в число независимых переменных включается оценка размера банка как логарифм чистых активов. При этом отмечается возможное нелинейное влияние этого показателя на вероятность дефолта. В качестве модели нелинейности предлагается использовать квадратичный полином. Аналогичное предположение делается относительно еще нескольких показателей. В результате оценивания модели авторы приходят к заключению о значимости нелинейной статистической связи с вероятностью отзыва лицензии для показателей, характеризующих размер банка, достаточность его капитала и рентабельность.

Описание используемых данных

О структуре выборки. Исходные данные для анализа представляют собой ежемесячную отчетность в открытом доступе 887 банков за период с 1 января 2013 г. по 1 декабря 2015 г. Этот временной интервал был выбран в связи со сменой руководства Банка России в середине 2013 г., повлекшей за собой существенное изменение характера задач, стоящих перед финансовой системой РФ. Была взята отчетность как действующих, в частности начавших работу, так и ликвидированных в этот период банков. Отметим, что некоторые из действующих в этот период банков позже - в 2016 и 2017 гг. также лишились лицензии.

Количество банков-резидентов в РФ постоянно сокращалось начиная с 1991 г. Для рассматриваемого промежутка времени по данным ЦБ РФ динамика этого показателя представлена в таблице 1.

Таблица 1

Динамика количества действующих банков и отзывов лицензий по данным ЦБ РФ

Дата	01.01.2013	01.01.2014	01.01.2015	01.12.2015
Количество действующих банков	956	923	834	740
Отозвано лицензий за год	29	73	88	

Важнейшей характеристикой деятельности банка является наличие лицензии на работу с физическими лицами. Пассивы банка образуют

собственные и привлеченные средства. Последние образованы клиентской базой и заимствованиями на межбанковском рынке. Статьи пассивов характеризуются срочностью и платностью; при этом платность статей возрастает с их срочностью. По возрастанию срочности и, соответственно, платности пассивы можно расположить следующим образом: средства на расчетных счетах, остатки на «лоро» счетах, межбанковские займы, депозиты предприятий и сберегательные счета граждан. Таким образом, при наличии соответствующей лицензии у банка возникает долгосрочный, хотя и не очень дешевый источник финансирования. В соответствии со статьей 6 ФЗ № 177 «О страховании вкладов физических лиц в банках Российской Федерации» все банки, имеющие лицензию на работу с физическими лицами, обязаны участвовать в системе страхования вкладов.

Представленная к анализу выборка банков на конец периода имеет структуру, представленную в таблице 2. Количество банков, получивших лицензию на работу с физическими лицами, незначительно менялось на протяжении выбранного периода времени. С учетом данных таблицы 1 можно отметить, что выбранная совокупность банков хотя и не исчерпывает их множество, но является представительной. По

остальным банкам в открытом доступе отчетности найти не удалось. Отметим, что из выборки было также удалено три банка, которые представили сомнительную отчетность.

Таблица 2

Структура выборки банков

Статус/АСВ	Не участвует	Участвует	Всего
Открыты	77	598	675
Закрыты	63	149	212
Всего	140	747	887

Различные критерии, в частности критерий Пирсона, а также критерий, основанный на коэффициенте тау-в Кендалла [16], позволяют отвергнуть гипотезу о независимости событий «отзыв лицензии» и «участие в системе страхования вкладов» с вероятностью ошибки первого рода 5%. Однако из данных таблицы 2 видно, что АСВ не является панацеей и немало банков, страхующих вклады, были закрыты в рассматриваемый период.

Анализ отчетности банков показывает, что медианы процентных долей основных заемных статей пассивов в общем их объеме (коэффициенты концентрации) существенно различаются для банков, участвующих и не участвующих в системе страхования вкладов. Результаты сравнения приведены в таблице 3.

Таблица 3

Структура заемного финансирования капитала банков. Банки, действующие на конец рассматриваемого периода

	Расчетные счета организаций	Счета «лоро»	Средства, привлеченные на межбанковском рынке	Средства юридических лиц	Средства физических лиц	Выпущенные облигации	Выпущенные векселя
Не участвует	0,152	0,008	0,126	0,108	0,001	0,081	0,026
Участвует	0,149	0,001	0,037	0,085	0,345	0,032	0,011
Все банки	0,150	0,001	0,040	0,086	0,330	0,033	0,012

Непараметрический критерий сравнения медиан [17] отвергает совпадение медиан с вероятностью ошибки первого рода 5% для всех статей пассивов, за исключением «Расчетных счетов организаций». Таким образом, структура финансирования банков, участвующих и не участвующих в системе страхования вкладов, существенно различается. Аналогичная ситуация наблюдается для структуры активов (см. таблицу 4).

На основании указанного выше теста можно заключить, что все медианы различны. Интересно, что для более чем 50% наблюдений (банк/месяц) отсутствуют долгосрочные инвестиции. Также банки, не участвующие в системе страхования вкладов, более интенсивно действуют на рынке межбанковских кредитов.

Приведенные выше результаты анализа позволяют предположить, что хозяйственная деятельность банков, работающих с физическими лицами, существенно отличается от хозяйственной деятельности банков, которые с ними не работают. И, следовательно, характеристики надежности этих групп банков могут сильно различаться - необходимо разделять их в анализе. В данном исследовании речь пойдет только об участниках системы страхования вкладов. Еще одним обстоятельством, которое необходимо учитывать, является причина отзыва лицензии. Можно выделить два случая - наличие и отсутствие причин, связанных с нарушением ФЗ № 115 «О противодействии легализации (отмыванию) доходов, полученных преступным путем, и фи-

Структура активов банков. Банки, действующие на конец рассматриваемого периода

	Денежные средства и драгоценные металлы	Корреспондентский счет в ЦБ РФ, включая резервы	Кредиты и депозиты банкам	Кредиты клиентам в сумме с резервами по ним	Приобретенные ценные бумаги	Долгосрочные инвестиции	Основные средства и нематериальные активы	Прочие активы
Не участвует	0,0262	0,0464	0,0985	0,5180	0,0012	0,0000	0,0065	0,0156
Участвует	0,0368	0,0395	0,0325	0,5508	0,0671	0,0000	0,0269	0,0236
Все банки	0,0360	0,0400	0,0360	0,5474	0,0596	0,0000	0,0242	0,0227

нансированию терроризма». Анализ отчетности показывает, что банки, закрытые по «криминальным» причинам, часто имеют весьма устойчивое с экономической точки зрения положение [14].

Таким образом, в центре внимания настоящего исследования находятся отзывы лицензий по «не криминальным причинам» у банков, которые имели лицензию на работу с физическими лицами.

Характеристики устойчивости банка

Одной из основных экономических причин утраты устойчивости банком является низкое качество активов и пассивов. К настоящему времени уже сложилась технология финансового анализа банка, в частности система показателей качества активов и пассивов [18]. Из приведенного выше обзора публикаций видно, что в большинстве исследований используются, помимо прочего, те или иные метрики, характеризующие достаточность собственных средств банка, обеспеченность его активами различной ликвидности, а также показатели, связанные с оценкой кредитной активности банка.

Остановимся на небольшой группе показателей, для значений которых установлены нормативы в инструкциях Банка России.

Характеристика достаточности собственного капитала банка (норматив Н1). Прежде всего следует иметь в виду, о каких суммах здесь идет речь. Согласно периодически обновляемому документу Центрального банка Российской Федерации «Алгоритм расчета показателей агрегированного балансового отчета 30 крупнейших банков Российской Федерации» по состоянию на 1 сентября 2015 г. фонды и прибыль для крупнейших банков РФ составляли не более 10% от общего объема пассивов. Таким образом, значимость этих сумм определяется в основном тем, что они являются собственностью хозяйствующего субъекта. Подробные инструкции о расчете данного показателя приведены в инструкции Центрального банка

Российской Федерации². Мы хотим обратить внимание на специфику его подсчета, которая подчеркивает его значимость для оценки устойчивости банка по сравнению с нередко используемым отношением величины собственного капитала к суммарной величине активов. В данном случае собственный капитал сопоставляется с возможными потерями, которые могут возникнуть при необходимости срочной реализации активов, и учитывается ликвидность активов. На основании значений этого показателя мы получаем информацию о том, насколько собственные средства банка покрывают эти потери. Поскольку очевидно, что активы формируются в основном за счет заемных средств, этот показатель позволяет оценить возможность банка выполнить свои обязательства при необходимости вернуть средства за счет реализации активов.

Характеристики достаточности активов по степени их ликвидности (нормативы Н2, Н3, Н4). В соответствии с определением устойчивости, приведенным выше, банк должен в любой момент времени выполнить все свои обязательства перед контрагентами, то есть вернуть заемные средства, которые формируют его пассивы. Пассивы можно классифицировать по степени срочности - краткосрочные обязательства, со сроком погашения в течение суток, среднесрочные, которые необходимо удовлетворить в течение 30 календарных дней и долгосрочные, со сроком погашения более года. Показатели нормативов Н2 и Н3, введенные Центральным банком Российской Федерации, позволяют оценить, насколько обязательства различной срочности обеспечены активами адекватной ликвидности, то есть какую долю обязательств можно удовлетворить за счет реализации той или иной группы активов. Нормативы, естественно, ограничивают эти отношения снизу. Что же касается показателя, связанного с долгосрочными активами (норматив Н4), то он позволяет сопоставить величину долгосрочных активов и долгосрочных источников финансирования,

² Инструкция об обязательных нормативах банков № 139 // Вестник Банка России. 2012. № 74. URL: <http://www.cbr.ru/publ/Vestnik/>.

включая собственный капитал. Таким образом, контролируется использование краткосрочных источников финансирования в формировании долгосрочных активов. Норматив, естественно, ограничивает указанное отношение сверху.

Включение в анализ прочих нормативов было ограничено наличием данных. Анализ показал, что на основании данных предоставляемой банками открытой отчетности можно включать в расчеты только показатели для нормативов Н7 и Н10.1. Прочие показатели имеют слишком много пропусков.

Показатель, связанный с нормативом Н7, позволяет контролировать обеспеченность крупных кредитных рисков собственным капиталом банка. К крупным рискам ЦБ РФ относит сумму кредитов, гарантий и поручительств в интересах одного клиента, которая превышает 5% собственного капитала банка. Показатель сопоставляет возможные совокупные потери по данным суммам с собственным капиталом банка.

И наконец, показатель, связанный с нормативом Н10.1, дает возможность сравнить сумму

кредитных рисков инсайдерам банка с его собственным капиталом. Под инсайдерами в этом случае понимаются физические лица, способные воздействовать на принятие решения о выдаче кредита банком. Таким образом, можно обнаружить, например, тенденцию вывода активов из банка.

Помимо указанных выше показателей, будем учитывать, есть ли в модели размер банка, поскольку во многих представленных в обзоре публикациях отмечается его значимое влияние на вероятность дефолта. В работах на тему устойчивости банков, не связанных с российским рынком банковских услуг, этому вопросу также уделяется много внимания [19]. Следуя уже сложившейся традиции, мы в качестве меры размера банка взяли логарифм его чистых активов. Под чистыми активами мы понимаем [18, 20] суммарный объем активов за вычетом расчетов с филиалами, транзитных счетов бюджета, расходов будущих периодов, текущих расходов, убытков, использования прибыли и выкупленных собственных акций. Описательные статистики этих показателей содержатся в таблице 5.

Таблица 5

Описательные статистики количественных показателей, включенных в модель

	H1_0D	H2D	H3D	H4D	H7D	H10_1D	Anet
Минимум	0,19	0,40	1,12	0,01	0,22	0,01	22509,19
Максимум	772,50	100000000,00	3720767,00	235,17	19072,80	119,48	18300000000,00
Среднее	25,24	46156,64	623,86	54,49	241,76	1,08	65600000,00
Нижняя квартиль	12,70	52,90	77,51	30,62	113,71	0,42	1392283,00
Медиана	17,44	75,77	100,31	54,26	213,11	0,91	3929229,00
Верхняя квартиль	29,16	113,27	139,18	78,60	341,92	1,63	14100000,00
Межквартильный размах	16,46	60,37	61,67	47,98	228,21	1,21	12700000,00
Стандартное отклонение	22,82	2136863,00	31804,44	30,24	222,40	1,16	620000000,00

Также в модель были внесены три фиктивные переменные, фиксирующие наступление важных для всей экономики или банковского рынка событий в рассматриваемый период: введение ключевой ставки в сентябре 2013 г., введение санкций против ряда участников российской экономики в сентябре 2014 г. и двукратное увеличение суммы покрытия в системе страхования вкладов вместе с распространением страхования на депозиты индивидуальных предпринимателей в декабре 2014 г.

Математические основы исследования

В подавляющем большинстве исследований, связанных с анализом устойчивости российских банков, используются различные варианты моделей конечного выбора. Их краткое описание

представлено в обзоре публикаций выше. Следует отметить, что практически все работы в этой области лишь в незначительной степени учитывают особенности панельной структуры данных, возникающей при использовании отчетности компаний за несколько периодов времени. Следует признать, что для корректного моделирования статистической взаимосвязи между вероятностями выбора альтернатив и значениями выбранных для исследования факторов в случае панельных данных может потребоваться специальный математический аппарат, подробно представленный, например, в [21].

Модели конечного выбора для панельных данных

Поясним кратко наше утверждение. Классическая модель конечного выбора основана на

представлении вероятности конкретного выбора с помощью функции распределения латентной переменной, «функции индекса» y_t :

$$(z_t = k) = P(y_t \in S_k); \quad S_k \subset R, \bigcup_k S_k = R. \quad (1)$$

Предполагается, что функция индекса связана с объясняющими факторами регрессионным соотношением вида

$$y_t = a_0 + G(x_{1,t}, \dots, x_{n,t}; a) + v_t. \quad (2)$$

В простейшем случае идет речь об обычной линейной регрессии. Предполагается, что выполняются основные гипотезы регрессионного анализа, и в частности v_t - независимые одинаково распределенные случайные величины.

Панельная структура данных предполагает повторные наблюдения над одними и теми же объектами в определенные моменты времени. Отметим, что эта структура данных хорошо вписывается в более широкую концепцию кластерных данных, которые формируются как наблюдения над объектами из нескольких фиксированных групп (кластеров). Например, хозяйствующие субъекты из нескольких фиксированных регионов. В случае панельных данных в качестве кластеров выступают представители панели, которые в различные моменты времени могут рассматриваться как разные объекты. Некоторым развитием моделей конечного выбора для этого случая служат модели с фиксированным эффектом, которые предполагают наличие индивидуальных констант для различных кластеров (элементов панели). Случайные составляющие по-прежнему считаются, как минимум, некоррелированными.

$$y_{i,t} = b_0 + b_i + G(x_{1,t}, \dots, x_{n,t}; a) + v_{i,t}. \quad (3)$$

И наконец, наиболее сложной ситуацией является наличие корреляции у случайной составляющей. Простейшей моделью учета этой особенности является модель со случайным эффектом, которая предполагает декомпозицию случайной составляющей на некоррелированную (независимую) компоненту и компоненту, специфичную для каждого кластера, которые не коррелируют между собой:

$$y_{i,t} = b_0 + G(x_{1,t}, \dots, x_{n,t}; a) + \mu_i + v_{i,t}. \quad (4)$$

Легко проверить, что в этом случае корреляционная матрица случайной составляющей будет иметь блочно-диагональный вид. Именно эта модель была использована в настоящем исследовании для решения поставленной задачи.

Отметим, что во многих приложениях приобрела популярность объединенная модель (модель со смешанным эффектом). Естественно, возможны расширения указанной модели, в которых наличие автокорреляции у случайной составляющей учитывается с помощью более сложных зависимостей. Подобное обсуждение особенностей спецификации и оценивания подобной модели выходит за рамки данной публикации и может быть найдено в [21]. Одна из основных проблем, как отмечено в этой работе, при игнорировании наличия корреляции у случайной составляющей заключается в возможной ошибочной оценке значимости регрессоров. Если корреляция действительно имеет место, то незначимые характеристики банков, которые могут меняться с течением времени, могут оцениваться как значимые.

Отметим, что учет панельной структуры данных позволяет по-другому взглянуть на такую важную проблему, возникающую при оценивании моделей классификации, в частности моделей конечного выбора, как несбалансированность данных. Известно, что если некоторые из выборов мало представлены в обучающем множестве, то оцененная (обученная) модель будет некорректно отображать истинное распределение вероятностей альтернатив. Модель может вообще «не заметить» некоторые редкие варианты выбора. Этой проблеме посвящено много исследований. В частности, о ней упоминается в работе [15]. Возможные риски, связанные с несбалансированными данными, и методы решения возникающих проблем подробно изложены в работе [22]. При этом обычно рассматриваются классически традиционные модели классификации, не учитывающие панельную (кластерную) структуру данных. Взглянем на проблему сбалансированности в контексте настоящего исследования. Используемые данные образованы повторными наблюдениями в различные моменты времени над одними и теми же объектами - банками. Модель бинарного выбора позволяет нам оценить вероятность дефолта банка в зависимости от значений объясняющих факторов, включенных в модель. Если фактор времени не входит явно в их число, то предполагается, что искомая вероятность не зависит явно от времени, в частности коэффициенты функции индекса не зависят от времени.

Различные варианты модели позволяют нам учесть специфику банка за счет индивидуального эффекта или корреляционной структуры данных. Но в любом случае, поскольку время

не учитывается в оценке вероятности дефолта, сбалансированность обучающего множества следует оценивать, соотнося количество дефолтов с количеством банков, а не общим количеством наблюдений - состояние конкретного банка в тот или иной месяц. В данном случае среди банков, получивших лицензию на работу с физическими лицами (747), было отозвано 149 лицензий. Нетрудно подсчитать, что закрытые банки составляют около 20% от исходного количества. В то же время общее количество наблюдений (включая пропущенные значения) за три года составляет 26892, и доля дефолтов существенно ниже - около 0,5%. Таким образом, степень несбалансированности обучающего множества существенно меняется в зависимости от способа ее оценки.

Обобщенные полиномы

Из приведенного выше обзора ясно, что во многих случаях авторы исследований предполагают, что по крайней мере, некоторые из объясняющих переменных имеют нелинейную статистическую взаимосвязь с вероятностью дефолта. При этом наиболее существенной особенностью этой взаимосвязи является наличие точки переключения направления влияния. Для проверки гипотезы о наличии связи такого рода используются квадратичные полиномы. Между тем класс функций, обладающих локальными экстремумами, безусловно, существенно более широк. Кроме того, во многих случаях исследователя может интересовать вопрос о свойствах статистической взаимосвязи при больших по абсолютной величине значениях независимой переменной. Не исключено, что в этом случае сила взаимосвязи постепенно уменьшается. Также является актуальным вопрос об эффекте изменения значения зависимой переменной на определенную величину в различных ситуациях. Линейная модель предполагает, что эффект всегда одинаков (коэффициент регрессии постоянен). И наконец, не всегда знание предметной области дает нам достаточно информации для правильной спецификации нелинейной модели. Нередко приходится полагаться на использование свойств имеющихся данных.

Обобщенные полиномы, введенные в работе [23], являются инструментом моделирования нелинейной статистической связи между зависимой переменной и объясняющими показателями. Детали спецификации и оценивания модели можно найти в работе [24]; их рассмотрение выходит за рамки настоящей публикации. Мы кратко

опишем частный случай этой модели, который был использован в исследовании. Обобщенные полиномы второго порядка позволяют учесть в модели особенности нелинейной связи, отмеченные выше.

Введем несколько обозначений. Обобщенной степенью x будем называть следующую функцию

$$x^{(p)} = \begin{cases} x^p, & p \neq 0 \\ \ln(x), & p = 0 \end{cases} \quad (5)$$

Обобщенным полиномом от x второго порядка мы будем называть функцию вида:

$$H(x; (a_1, a_2), (p_1, p_2)) = \begin{cases} a_1 x^{(p_1)} + a_2 x^{(p_2)}, & p_1 \neq p_2 \\ a_1 x^{(p)} + a_2 x^{(p)} \ln(x), & p_1 = p_2 = p \end{cases} \quad (6)$$

В дальнейшем мы будем называть эту функцию функцией индекса переменной x .

Таким образом, для нескольких переменных-нелинейная модель их статистической взаимосвязи с зависимой переменной имеет (опуская параметры) вид:

$$y_i = \sum_k H_k(x_{k,i}; (a_{1,k}, a_{2,k}), (p_{1,k}, p_{2,k})) + v_i \quad (7)$$

Модель легко обобщается на панельные данные. Важно, что значения как коэффициентов, так и степеней определяются по обучающему множеству в результате использования оптимизационной процедуры. Фактически в указанных выше работах был введен класс нелинейных моделей, в котором выбирается экземпляр, наиболее адекватный данным. Модели такого вида получили широкое распространение в медицине и биологии, однако нам не приходилось сталкиваться с их использованием при решении эконометрических задач. Также отметим, что модель линейной регрессии является частным случаем моделей этого типа.

Описание результатов моделирования

Перейдем к описанию результатов моделирования. Спецификация модели, выбранная для решения поставленной в исследовании задачи, позволяет утверждать, что вероятность отзыва лицензии i -го банка может быть представлена в виде:

$$p(z_i = 1 | x_1, \dots, x_n; a, p, \alpha_i) = f(y(x_1, \dots, x_n; a, p, \alpha_i)); f(z) = \frac{1}{1 + e^{-z}} \quad (8)$$

$$y(x_1, \dots, x_n; a, p, \alpha_i) = \sum_{k=1}^n H_k(x_{k,i}; (a_{1,k}, a_{2,k}), (p_{1,k}, p_{2,k})) + \alpha_i$$

Здесь предполагается, что i -й банк представляет гипотетический кластер банков, который характеризуется постоянной величиной случайной константы - α_i . Для конкретных вычислений параметры функций индекса объясняющих факторов - коэффициенты $(a_{1,k}, a_{2,k})$ и степени $(p_{1,k}, p_{2,k})$, а также случайная константа заменяются их оценками.

Аналогично классической модели бинарного выбора в случае логистической регрессии с обобщенными полиномами направление и силу влияния фактора на вероятность отзыва лицензии при заданных значениях всех факторов можно оценить по маргинальному эффекту данного фактора (значению производной вероятности по фактору; часть аргументов в выражении для краткости опущена):

$$ME_k^{fp} = \frac{\partial}{\partial x_k} p(z_i = 1) = f(y)(1 - f(y)) \frac{\partial}{\partial x_k} H_k(x_k). \quad (9)$$

Для классической логистической регрессии маргинальный эффект имеет вид:

$$ME_k^0 = f(y)(1 - f(y)) \alpha_k,$$

где α_k - коэффициент при данном факторе в функции индекса.

Таким образом, направление влияния фактора на вероятность отзыва лицензии определяется участками монотонности функции индекса этого фактора. Также следует обращать внимание на поведение функции индекса при увеличении значений фактора по абсолютной величине. Наличие горизонтальных асимптот говорит о том, что в этом случае сила влияния фактора постепенно убывает.

Проанализируем поведение функций индекса объясняющих факторов в решаемой задаче. В таблице 6 приведены результаты оценивания классической логит-модели и логит-модели с обобщенными полиномами. Для оценивания второй модели была использована модификация алгоритма MFP [2] без удаления незначимых переменных на каждом шаге алгоритма.

Таблица 6

Результаты оценивания моделей

	Классическая модель					Модель с обобщенными полиномами				
	scale	p1	p2	a1	a2	scale	p1	p2	a1	a2
H1	1	1		-0,0430* (0,0175)		100	-1	-1	1,967*** (0,366)	0,604*** (0,115)
H2	1	1		0,00234 (0,00143)		1000	-1	-1	0,116*** (0,0229)	0,0165*** (0,00334)
H3	1	1		-0,0290*** (0,00478)		1000	-0,5	-0,5	-1,553 (1,885)	-0,536 (0,423)
H4	1	1		-0,0130** (0,00428)		100	2	3	-6,459*** (1,897)	5,052** (1,712)
H7	1	1		0,00121*** (0,00023)		10000	1	1	155,8*** (32,37)	59,45*** (13,84)
H10_1	1	1		-0,0315 (-0,043)		10	2	2	0,46 (1,609)	-3,759 (4,132)
Anet	1	1		-1,45e-08* (6,37e-9)		1,0E+10	-1	3	0,000111*** (2,5e-5)	-352734,8 (243626,8)
keyrate				2,329** (0,731)					2,097** (0,735)	
sanction				0,163 (0,402)					-0,0146 (0,481)	
maxcomp				0,536 (0,404)					0,79 (0,486)	
_cons				-4,111*** (0,873)					-12,07*** (3,062)	
lnsig2u				-10,97 (18,32)					-5,939 (18,74)	
N				21915					21915	
AIC				896,2					763,8	

Примечание: *, **, *** - значимость оценки на 5%-, 1%-, 0,1%-ном уровне.

В столбцах «scale» приведены коэффициенты масштабирования показателей. В случае

классической модели они равны 1 (масштабирование не производится), в случае модели с

обобщенными полиномами коэффициенты масштабирования определялись автоматически при реализации алгоритма оценивания. В столбцах p_1, p_2 приведены обобщенные степени показателя. Для модели с обобщенными полиномами они выбираются автоматически. В столбцах a_1, a_2 приведены оценки коэффициентов обобщенных полиномов, в круглых скобках - стандартные ошибки.

Функция индекса для показателя 'H1_0' имеет вид:

$$H(H1) = a_1 \left(\frac{H1}{scale} \right)^{-1} + a_2 \left(\frac{H1}{scale} \right)^{-1} \ln \left(\frac{H1}{scale} \right). \quad (10)$$

Оценки обоих коэффициентов значимы и были использованы при подсчете значений функции индекса.

Как следует из работы [24], эта функция имеет максимум в точке $H1^* \approx 10,4\%$, что достаточно близко к нормативу ЦБ РФ для данного показателя ($H1 > 10\%$). Кроме того, функция индекса стремится к нулю при увеличении значений показателя. Таким образом, можно заключить, что при фиксированных значениях остальных метрик значения показателя $H1$, меньшие точки максимума, статистически положительно связаны с вероятностью отзыва лицензии, а значения, большие этой величины, статистически отрицательно связаны с этой вероятностью. Также значительный рост значений этого показателя сводит на нет его статистическую связь с вероятностью отзыва лицензии, то есть очень большие его значения можно не принимать во внимание.

Функция индекса для показателя $H2$ (мгновенная ликвидность) имеет аналогичный вид. Максимум функции достигается в точке $H2^* \approx 2,4\%$. Минимально допустимое значение для этого показателя по мнению ЦБ РФ существенно выше - 15%. Расчеты показывают, что на практике допустимы существенно более низкие значения. При больших значениях показателя его связь с вероятностью отзыва лицензии теряется. Отметим, что в классической модели бинарного выбора оценки коэффициента для данного показателя в функции индекса модели незначима.

Ситуация с показателем среднесрочной ликвидности менее однозначна. Оценки обоих коэффициентов в модели с обобщенными полиномами незначимы. Формально есть основания считать, что взаимосвязь с вероятностью дефолта отсутствует. Однако оценка коэффициента для этой метрики в классической модели значима.

В связи с этим мы оставили данный показатель в модели. Функция его индекса имеет вид:

$$H(H3) = a_1 \left(\frac{H3}{scale} \right)^{-0.5} + a_2 \left(\frac{H3}{scale} \right)^{-0.5} \ln \left(\frac{H3}{scale} \right). \quad (11)$$

Функция этого индекса имеет минимум в точке $H3^* \approx 408\%$. Это означает, что увеличение данного показателя до точки минимума статистически связано с уменьшением вероятности дефолта. Доля измерений, превышающих указанное значение, не очень велика. Как следует из таблицы 3, 75% измерений (банк-месяц) не превышают 139,18%. Более того, отметим, что не менее 95% измерений имеют значение показателя, не превышающее значение точки минимума. Таким образом, при фиксированных значениях прочих показателей в большинстве случаев увеличение среднесрочной ликвидности банка благотворно сказывается на его устойчивости.

Индекс показателя долгосрочной ликвидности имеет вид:

$$H(H4) = a_1 \left(\frac{H4}{scale} \right)^2 + a_2 \left(\frac{H4}{scale} \right)^3. \quad (12)$$

Функция имеет минимум в точке $H4^* \approx 85,2\%$. Эта величина существенно меньше норматива ЦБ РФ (<120%), однако характер ограничения сохраняется. Большие значения показателя статистически связаны с ростом вероятности дефолта при постоянных значениях прочих показателей.

Рассмотрим показатель обеспеченности крупных кредитных рисков $H7$. Его функция индекса имеет вид:

$$H(H7) = a_1 \left(\frac{H7}{scale} \right) + a_2 \left(\frac{H7}{scale} \right) \ln \left(\frac{H7}{scale} \right). \quad (13)$$

Она имеет точку минимума $H7^* \approx 26,8\%$, что существенно меньше, чем норматив ЦБ РФ (<800%). Однако характер ограничения сохраняется. Отметим, что оценка коэффициента для данного показателя в классической модели также значима и положительна. Как следует из таблицы 3, более 25% банков имеют значение данного показателя превышающее значение точки минимума.

Индекс показателя крупных кредитных рисков инсайдером банка ($H10_1$) имеет незначимые оценки коэффициентов как в классической модели, так и в модели с обобщенными полиномами. В связи с этим мы не будем анализировать его функцию индекса.

Что касается размеров банка, то функция индекса чистых активов имеет вид:

$$H(Anet) = a_1 \left(\frac{Anet}{scale} \right)^{-1} + a_2 \left(\frac{Anet}{scale} \right)^3. \quad (14)$$

Это убывающая выпуклая вверх функция, у которой нет экстремумов. Таким образом, увеличение размера банка положительно статистически связано с вероятностью дефолта.

Что же касается фиктивных переменных, то в обеих моделях значимую оценку коэффициента имеет только переменная, связанная с введением ключевой ставки. Это событие повысило вероятность отзыва лицензии у всех банков. По времени оно практически совпадает с началом реформ банковского рынка.

Заключение

Приведенные выше результаты исследования позволяют заключить, что использование автоматического определения структуры модели на основе свойств данных в указанном случае обобщенных полиномов может существенно изменить наше представление о статистической взаимосвязи между вероятностью интересующего нас события (отзыва лицензии) и факторами, характеризующими изучаемый объект и среду, в которой он функционирует.

Мы видим, что для ряда факторов классическая модель бинарного выбора и модель с обобщенными полиномами по-разному оценивают значимость одних и тех же независимых переменных для формирования вероятности отзыва лицензии. Справедливо задать вопрос - какая

модель более адекватна свойствам данных. Обратимся к классическим метрикам оценки качества моделей конечного выбора для сравнения двух указанных выше вариантов. Отметим, что в данном исследовании не ставилась задача построения наилучшей в некотором смысле прогностической модели. На первом плане было качество описания данных.

Значения статистики критерия Акаике (AIC), приведенные в таблице 6, говорят о том, что с точки зрения соотношения сложности и описательной точности модели модель с обобщенными полиномами лучше. Аналогичный результат дает использование Байесовского информационного критерия, для классической модели значение статистики - 992,2, для модели с обобщенными полиномами - 915,7. Однако, безусловно, это далеко не все, что требуется от моделей данного типа.

Рассмотрим качество интерполяции событий для обеих моделей. Выберем в качестве порогового значения вероятности дефолта $p^* = 0,5$ и будем считать, что модель идентифицирует наблюдение «банк-месяц» как соответствующее отзыву лицензии, если вероятность дефолта, вычисленная моделью, больше этой величины. Рассмотрим таблицы сопряженности истинных и интерполированных состояний банков для обоих типов моделей. В данном анализе участвуют все имеющиеся измерения, а не только те, что использовались для оценивания коэффициентов моделей.

Таблица 7

Таблицы сопряженности для двух типов моделей

	Классическая модель			Модель с обобщенными полиномами		
	Открыт	Закрыт	Всего	Открыт	Закрыт	Всего
Истинные/интерполированные						
Открыт	25238	6482	31720	21584	10136	31720
Закрыт	144	68	212	62	150	212
Всего	25382	6550	31932	25382	6550	31932
Чувствительность (Sensitivity)	0,32			0,7		
Специфичность (Specificity)	0,8			0,68		

Видно, что для модели с обобщенными полиномами чувствительность, то есть доля правильно интерполированных дефолтов, более чем в два раза больше чувствительности классической модели. Однако, естественно, специфичность, то есть доля правильно интерполированных отсутствий дефолта, у этой модели меньше.

Следует иметь в виду, что данный анализ проведен для фиксированного порогового значения вероятности дефолта. Варьирование порога от

нуля до единицы приводит нас к так называемой ROC- кривой, графику чувствительности относительно величины «1 - специфичность», то есть доли неверно интерполированных дефолтов. Как отмечено в работе [21], площадь под этой кривой характеризует способность модели различать на основе значений выбранных показателей наблюдения, обладающие и не обладающие интересующим исследователя признаком, в данном случае наблюдения, для которых произошел или

не произошел дефолт. Как показывают расчеты, площадь под ROC-кривой для классической модели составляет примерно 0,7, а для модели с обобщенными полиномами - примерно 0,9, что свидетельствует в пользу второй модели.

И, наконец, обратимся к критерию Hosmer-Lemeshow [21], который позволяет проанализировать, насколько вероятность события, в данном случае дефолта, согласуется с наблюдаемой частотой его появления. Для классической модели статистика критерия равна 3863,33 и согласование отвергается с вероятностью ошибки первого рода менее 1%. Для модели с обобщенными полиномами статистика критерия равна 6,24 и согласование не отвергается (p -уровень равен 0,62).

Проведенный выше анализ качества моделей показывает, что по ряду показателей модель с обобщенными полиномами превосходит классическую модель. Однако следует отметить, что многие из приведенных выше критериев были разработаны для случая, когда данные не имеют панельной структуры. В частности, везде идет речь об оценке вероятности отзыва лицензии у конкретного банка в конкретный месяц. В то же время, как отмечалось выше, в рамках дистанционного мониторинга, как правило, на первом плане стоит вопрос об устойчивости конкретного банка безотносительно момента времени. В связи с этим к преимуществу модели с обобщенными полиномами следует относиться с осторожностью и использовать ее совместно с классической моделью.

Отметим также результаты проверки наличия автокорреляции у случайной составляющей функции индекса в использованных моделях бинарного выбора. Напомним, что в связи с предположением ее наличия использовались модели со случайным эффектом. Результаты оценивания показывают, что, по крайней мере, для этого способа учета корреляционной структуры гипотеза об отсутствии автокорреляции не отвергается. Таким образом, либо автокорреляция отсутствует и можно пользоваться более простыми вариантами модели, либо она имеет более сложный вид.

Литература

1. Райзберг Б.А., Лозовский Л.Ш., Стародубцева Е.Б. Современный экономический словарь. 6-е изд., перераб. и доп. М.: ИНФРА-М, 2011. 203 с.
2. Тавасиев А.М., Бычков В.П., Москвин В.А. Банковское дело: базовые операции для клиентов. М.: Финансы и статистика, 2005. 304 с.

3. Berk R.A. Statistical Learning from a Regression Perspective. Springer, 2008. 369 p.

4. Sauerbrei W., Royston P., Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building // Statistics in medicine. 2007. No. 26. P. 5512-5528. Published online in Wiley Inter Science. URL: www.interscience.wiley.com.

5. Sahajwala R., Van den Bergh P. Supervisory Risk Assessment and Early Warning Systems//BCBS Working Paper. 2000. No. 4. Retrieved from. URL: http://www.bis.org/publ/bcbs_wp4.pdf.

6. Guidance on the application of the Core Principles for Effective Banking Supervision to the regulation and supervision of institutions relevant to financial inclusion. BCBS, 2016.

7. Помазанов М.В. Управление кредитным риском в банке: подход внутренних рейтингов: практическое пособие для магистратуры. Под науч. ред. Г.И. Пенникаса. М.: Юрайт, 2017. 265 с.

8. Тотьянина К. Обзор моделей вероятности дефолта // Управление финансовыми рисками. 2011. № 1 (25).

9. Bluhm C., Overbeck L., Wagner C. An introduction to credit risk modeling // Chapman & Hall/CRC, Boca Raton, 2003. 285 p.

10. Martin D. Early warning of bank failure: A logit regression approach // Journal of Banking and Finance. 1977. Vol. 1 (3). P. 249-276.

11. Головкин Е.Л., Сидоров В.Г., Пересецкий А.А., Карминский А.М., ван Султ А.Г.О. Анализ рейтингов российских банков // Препринт #2002/033. М.: Российская экономическая школа, 2002. 37 с.

12. Головань С.В., Карминский А.М., Копылов А.В., Пересецкий А.А. Модели вероятности дефолта российских банков. I. Предварительное разбиение банков на кластеры. Препринт #2003/039. М.: Российская экономическая школа, 2003.

13. Sharma S., Shebalkov M., Yukhanaev A. Evaluating banks performance using key financial indicators - a quantitative modeling of Russian banks // The Journal of Developing Areas. 2016. Vol. 50. No. 1. P. 425-453.

14. Пересецкий А.А. Модели причин отзыва лицензий российских банков. Влияние неучтенных факторов // Прикладная эконометрика. 2013. № 30 (2). С. 49-64.

15. Карминский А.М., Костров А.В. Моделирование вероятности дефолта российских банков: расширенные возможности // Журнал Новой экономической ассоциации. 2013. № 1 (17). С. 64-86.

16. Аптон Г. Анализ таблиц сопряженности. М.: Финансы и статистика, 1982. 143 с.

17. Snedecor G.W., Cochran W.G. Statistical Methods. 8th ed. // Ames, IA: Iowa State University Press, 1989.

18. Смирнов А.В. Анализ финансового состояния коммерческих банков. М., 2007. 225 с. URL: https://www.biznesbooks.com/components/com_jshopping/

files/demo_products/smirnov-a-v-analiz-finansovogo-sostoyaniya-kommercheskikh-bankov.pdf.

19. **Steinbacher M., Steinbacher M.** How banks' capital ratio and size affect the stability of the banking system: asimulation-based study // *Journal of Credit Risk*. 2014. No. 11 (1). P. 59-92.

20. **Жарковская Е. П.** Финансовый анализ деятельности коммерческого банка. М.: ОМЕГА-Л, 2011.

21. **Hosmer D.W., Lemeshow S., Sturdivant R.X.** *Applied Logistic Regression*. 3rd ed. John Wiley & Sons. 2013.

22. **He H., Edwardo A.** Learning from Imbalanced Data // *IEEE Transactions on Knowledge and Data Engineering*. 2002. Vol. 21. No. 9. P. 1263-1284.

23. **Royston P., Altman D.G.** Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling // *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. 1994. Vol. 43. No. 3. P. 429-467.

24. **Royston P., Sauerbrei W.** Multivariable model-building : a pragmatic approach to regression analysis based on fractional polynomials for continuous variables // John Wiley & Sons. 2008.

MODELING SUSTAINABILITY OF THE RUSSIAN BANKS AMID BANKING SYSTEM REFORM

Konstantin L. Polyakov

Author affiliation: National Research University Higher School of Economics (Moscow, Russia). E-mail: polyakov.kl@hse.ru.

Marina V. Polyakova

Author affiliation: National Research University Higher School of Economics (Moscow, Russia). E-mail: mpolyakova@hse.ru.

This article presents results of mathematical and statistical modeling research of sustainability of banks' functioning that is relevant in light of banking system reforms. Basic concept for these constructions lies in the understanding of stability (or reliability) of business operation, financial state of which in normal circumstances ensures the fulfillment of all its obligations to the employees, other organizations and the State due to sufficient income and the matching of cost with revenue. Some of the most significant for the analysis of banks' stability (reliability) indicators are available for all users (individuals as well as organizations) in the form of monthly financial reports submitted by the majority of banks, which are uploaded to the official site of the Bank of Russia and are duly and timely updated. In the authors' opinion these indicators include some for which the Bank of Russia defines normative values. Although, monitoring the degree of consistency between actual indicators and regulations is far from being an ideal instrument to analyze bank reliability (i.e. to determine the situation that does not lead to license revocation), values of these indicators in the authors' opinion are sufficient characteristics to establish financial health of a credit institution.

This article studies the nature of statistical relationship between trends of quantity values of the most vital aspects of banks' activities and the probability of license revocation. The specific feature of models' construction is the automatic choice of the functional entry form for bank characteristics. For this purpose, the authors use the generalized polynomials, which makes it possible to select model specification most comparable to data attributes.

The study was conducted on the basis of the public records of 887 banks for the period from 01.01.2013 to 01.12.2015. For this analysis, the authors used reports from already operating banks, in particular those that have just started operation, as well as banks that were liquidated in that period (some of the banks that operated during the research period later, in 2016 and 2017, also lost their licenses). The results of the model evaluation demonstrated a high degree of comparability of the entry form for bank characteristics with their economic substance and the Bank's of Russia regulations. Quality comparison between the classical binary choice model for panel data and a model based on generalized polynomials shows a distinct advantage of the latter.

Keywords: bank stability, license revocation, banks reporting, banking statistics, binary choice model, panel data, generalized polynomials.

JEL: C33, C51, G01, G18, G21, G22, G28.

References

1. **Raizberg B.A., Lozovskii L.Sh., Starodubtseva E.B.** *Sovremenniy ekonomicheskii slovar'*. 6-e izd., pererab. i dop. [Modern economic dictionary. 6th edition reviewed and updated]. Moscow, INFRA-M Publ., 2011. 203 p. (In Russ.).

2. **Tavasiev A.M., Bychkov V.P., Moskvina V.A.** *Bankovskoe delo: bazovye operatsii dlya klientov* [Banking: basic operations for clients]. Moscow, Finansy i statistika Publ., 2005. 304 p. (In Russ.).

3. Berk R.A. Statistical Learning from a Regression Perspective. Springer, 2008. 369 p.
4. Sauerbrei W., Royston P., Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Statistics in medicine*, 2007, no. 26, pp. 5512-5528. Published online in Wiley Inter Science. Available at: www.interscience.wiley.com.
5. Sahajwala R., Van den Bergh P. Supervisory Risk Assessment and Early Warning Systems. *BCBS Working Paper*, 2000, no. 4. Retrieved from: http://www.bis.org/publ/bcbs_wp4.pdf.
6. Guidance on the application of the Core Principles for Effective Banking Supervision to the regulation and supervision of institutions relevant to financial inclusion. BCBS, 2016.
7. Pomazanov M.V., Penikas G.I. (Ed.). Upravlenie kreditnym riskom v banke: podkhod vnutrennikh reitingov: prakticheskoe posobie dlya magistratury [Credit risk management in the bank: internal rating-based approach. Practical manual for Master's Degree course]. Moscow, Yurait Publ., 2017. 265 p. (In Russ.).
8. Tot'myanina K. Obzor modelei veroyatnosti defolta [Review of models of default probability]. *Upravlenie finansovymi riskami*, 2011, no. 1 (25). (In Russ.).
9. Bluhm C., Overbeck L., Wagner C. An introduction to credit risk modeling. Chapman & Hall/CRC, Boca Raton, 2003. 285 p.
10. Martin D. Early warning of bank failure: A logit regression approach. *Journal of Banking and Finance*, 1977, vol. 1 (3), pp. 249-276.
11. Golovko E.L., Sidorov V.G., Peresetskii A.A., Karminskii A.M., van Sust A.G.O. Analiz reitingov rossiiskikh bankov. Preprint #2002/033 [Analysis of ratings of Russian banks. Preprint #2002/033]. Moscow, New Economic School, 2002. 37 p. (In Russ.).
12. Golovan S.V., Karminskii A.M., Kopylov A.V., Peresetskii A.A. Modeli veroyatnosti defolta rossiiskikh bankov. I. Predvaritel'noe razbienie bankov na klasteri. Preprint #2003/039 [Models of probability of default of Russian banks. I. Preliminary breaking up of banks into clusters. Preprint #2003/039]. Moscow, New Economic School, 2003. (In Russ.).
13. Sharma S., Shebalkov M., Yukhanaev A. Evaluating banks performance using key financial indicators - quantitative modeling of Russian banks. *The Journal of Developing Areas*, 2016, vol. 50, no. 1, pp. 425-453.
14. Peresetsky A.A. Modeli prichin otzyva litsenzii rossiiskikh bankov. Vliyanie neuchtennykh faktorov [Modeling reasons for Russian bank license withdrawal: Unaccounted factors]. *Applied Econometrics*, 2013, no. 30 (2), pp. 49-64. (In Russ.).
15. Karminskii A.M., Kostrov A.V. Modelirovanie veroyatnosti defoltarossiiskikh bankov: rasshirennye vozmozhnosti [Modeling the Default Probabilities of Russian Banks: Extended Abilities]. *The Journal of the New Economic Association*, 2013, no. 1 (17), pp. 64-86. (In Russ.).
16. Apton G. Analiz tablits sopryazhennosti [Analysis of cross tables]. Moscow, Finansy i statistika Publ., 1982. 143 p. (In Russ.).
17. Snedecor G.W., Cochran W.G. Statistical Methods. 8th ed. Ames, IA: Iowa State University Press, 1989.
18. Smirnov A.V. Analiz finansovogo sostoyaniya kommercheskikh bankov [Analysis of the financial condition of commercial banks]. Moscow, 2007. 225 p. (In Russ.). Available at: https://www.biznesbooks.com/components/com_jshopping/files/demo_products/smirnov-a-v-analiz-finansovogo-sostoyaniya-kommercheskikh-bankov.pdf.
19. Steinbacher M., Steinbacher M. How banks' capital ratio and size affect the stability of the banking system: a simulation-based study. *Journal of Credit Risk*, 2014, no. 11 (1), pp. 59-92.
20. Zharkovskaya E. P. Finansovyi analiz deyatel'nosti kommercheskogo banka [Financial analysis of the activities of a commercial bank]. Moscow, OMEGA-L Publ., 2011. (In Russ.).
21. Hosmer D.W., Lemeshow S., Sturdivant R.X. Applied Logistic Regression. 3rd ed. John Wiley & Sons, 2013.
22. He H., Edwardo A. Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*, 2002, vol. 21, no. 9, pp. 1263-1284.
23. Royston P., Altman D.G. Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 1994, vol. 43, no. 3, pp. 429-467.
24. Royston P., Sauerbrei W. Multivariable model-building : a pragmatic approach to regression analysis based on fractional polynomials for continuous variables. John Wiley & Sons, 2008.