

ВЫСШАЯ ШКОЛА ЭКОНОМИКИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

А.М. Карминский, А.В. Костров, Т.Н. Мурзенков

**МОДЕЛИРОВАНИЕ ВЕРОЯТНОСТИ
ДЕФОЛТА РОССИЙСКИХ БАНКОВ
С ИСПОЛЬЗОВАНИЕМ
ЭКОНОМЕТРИЧЕСКИХ МЕТОДОВ**

Препринт WP7/2012/04

Серия WP7

Математические методы
анализа решений в экономике,
бизнесе и политике

Москва
2012

УДК 330.43
ББК 65.в6
К24

Редакторы серии WP7
«Математические методы анализа решений в экономике,
бизнесе и политике»

Ф.Т. Алескеров, В.В. Подиновский, Б.Г. Миркин

К24

Карминский, А. М., Костров, А. В., Мурзенков, Т. Н. Моделирование вероятности дефолта российских банков с использованием эконометрических методов : препринт WP7/2012/04 [Текст] / А. М. Карминский, А. В. Костров, Т. Н. Мурзенков; Нац. исслед. ун-т «Высшая школа экономики». – М. : Изд. дом Высшей школы экономики, 2012. – 64 с. – 50 экз.

Совершенствование моделей вероятности дефолта банков является одним из перспективных направлений риск-менеджмента, предусмотренных новым Базельским соглашением в рамках IRB-подхода. В данном исследовании особое внимание уделяется:

- расширению горизонта эмпирического исследования за счет использования российской банковской статистики за период с 1998 г. по 2011 г.;
- оценке влияния макроэкономических и институциональных факторов на вероятность дефолта банка;
- тестированию качества построенной модели.

Кроме того, в работе учитывается влияние рыночной структуры банковского сектора РФ и нелинейностей по объясняющим переменным на вероятность дефолта банка, а также уделено внимание тестированию достоверности моделей.

Проведен сравнительный анализ эконометрических моделей вероятности дефолта с альтернативными. В качестве базовой была использована логистическая регрессия с квазипанельной структурой данных. В результате мы пришли к следующим выводам:

- присутствует квадратичная зависимость между достаточностью капитала банка и вероятностью его дефолта;
- существует отрицательная зависимость между уровнем конкуренции, характеризующим индексом Лернера, и вероятностью дефолта банка;
- учет макроэкономических и институциональных переменных, как и фактора времени, существенно улучшает качество модели.

Результаты моделирования представляют потенциальный интерес не только для регулятора, но и для коммерческих банков в рамках задач риск-менеджмента.

Ключевые слова: вероятность дефолта, банки, Россия, риск-менеджмент, внутренние рейтинги, IRB, Базель II.

JEL: G21, G24, G32.

УДК 330.43

ББК 65.в6

Работа выполнена в рамках научно-учебной группы «Моделирование дефолтов кредитных организаций», созданной по программе «Научный фонд НИУ ВШЭ» (Учитель – Ученики) по проекту № 05-0030 «Анализ и моделирование дефолтов кредитных организаций», 2012 г.

**Препринты Национального исследовательского университета
«Высшая школа экономики» размещаются по адресу: <http://www.hse.ru/org/hse/wp>**

© Карминский А. М., 2012

© Костров А. В., 2012

© Мурзенков Т. Н., 2012

© Оформление. Издательский дом
Высшей школы экономики, 2012

Содержание

Введение	4
1. Предпосылки построения моделей вероятности дефолта для российских банков	6
1.1. Российская банковская система и проблема устойчивости банков	6
1.2. Моделирование вероятности дефолта: обзор литературы	12
2. Формирование базы данных для эмпирического моделирования вероятности дефолта банка	18
2.1. Понятие и классификация видов дефолта	18
2.2. Источники национальной банковской статистики	19
2.4. Описание модели бинарного выбора: logit-модель.....	23
3. Построение модели бинарного выбора для вероятности дефолта банка	24
3.1. Первичный отбор и очистка объясняющих переменных	24
3.2. Проблема несбалансированности данных	31
3.3. Выбор величины лага по объясняющим переменным	32
3.4. Анализ нелинейностей по финансовым переменным.....	35
3.5. Анализ фактора времени.....	37
3.6. Учет макроэкономических и институциональных факторов	38
3.7. Тестирование качества построенной модели	45
3.8. Интерпретация результатов	47
3.9. Тестирование предсказательной силы модели. Прогноз вне выборки.....	53
4. Сравнение альтернативных моделей вероятности дефолта	54
4.1. Логистическая модель бинарного выбора с использованием панельных данных.....	54
4.2. Нейронная сеть для прогнозирования вероятности дефолта банка.....	55
4.3. Модель с использованием в качестве объясняющей переменной Z-индекса	56
Заключение	56
Литература	57
Приложения	60

Введение

В данном исследовании рассмотрены особенности моделирования вероятности дефолта банка в условиях российской действительности при помощи логистической модели бинарного выбора. На основе национальной банковской статистики, макроэкономических и институциональных данных за 1998–2011 гг. в работе построен ряд моделей вероятности дефолта (моделей раннего предупреждения) для российского банковского сектора.

Моделирование проводилось при помощи стандартной логистической регрессии, которая позволяет детально оценить влияние различных факторов на предмет исследования – вероятность дефолта банка, а также предоставляет возможность качественной интерпретации полученных результатов.

В предшествующих исследованиях акценты были сделаны преимущественно на построении моделей, которые не имели отношения к России, либо акцентировались периоды кризиса, преимущественно 1998 г. В то же время наиболее приемлемым выглядит использование таких моделей в относительно стабильных условиях, когда зарождаются предпосылки неустойчивости как отдельных банков, так и банковской системы в целом.

Поэтому мы использовали квазипанельные данные за сравнительно большой период времени – с 1998 по 2011 г. Кроме того, в данном исследовании внимание было сосредоточено на таких факторах, как время, макроэкономические и институциональные характеристики операционной банковской среды, а также на возможном влиянии нелинейностей по финансовым переменным на вероятность дефолта банка.

Можно надеяться, что динамика вероятностей дефолта банков подскажет Банку России момент времени и условия для изменения требований к капиталу, помогая регулятору бороться с проблемой процикличности.

Начиная эту работу, мы предполагаем, что модель вероятности дефолта будет востребована тремя основными агентами: Банком России как регулятором, коммерческими банками и их контрагентами. При помощи модели Банк России сможет выявлять наиболее уязвимые банки (группу риска) и при необходимости своевременно принимать меры для их финансового оздоровления, учитывать слабость конкретной группы банков при реализации своей политики. Для коммерческих

банков (КБ) наблюдение как за динамикой своей вероятности дефолта, так и за соответствующими характеристиками контрагентов, позволит получить независимую оценку устойчивости и перспектив развития банка, а также количественно оценивать влияние своих действий на риски, связанные с банковской деятельностью. Модель может оказаться полезной для контрагентов банка, помогая им лучше осознавать риски вложения финансовых средств в конкретный банк.

Цель данной работы – предложить адекватную модель, которая поможет пользователю предсказывать дефолты банков. Такая модель с высокой вероятностью будет востребована национальным банковским сектором.

В соответствии с поставленной целью необходимо:

- обеспечить сбор необходимых для исследования финансовых данных об операционной деятельности банков и информации о внешней среде, с которой им приходится взаимодействовать;
- осуществить отбор финансовых, макроэкономических и институциональных факторов, влияющих на вероятность дефолта кредитной организации, определение характера этого влияния;
- определить спецификации моделей вероятности дефолта банка;
- провести сравнение полученных моделей, протестировать их качество, предсказательную силу;
- сравнить возможности панельных моделей, построенных по усеченным выборкам, с ранее полученными моделями.

В данной работе будет использована модель бинарного выбора для оценки вероятности дефолта, в частности, логистическая регрессия. С ее помощью естественно оценивается влияние на вероятность дефолта каждого из используемых объясняющих факторов, что позволяет рассчитать соответствующие предельные эффекты. В последнее время логистическая модель с панельной структурой данных получила широкое распространение среди исследователей, ею воспользуемся в данной работе и мы.

Исследование имеет следующую структуру. В первом разделе представлен обзор российской банковской системы и литературы по моделированию дефолта банков. Во втором разделе описаны источники построенной базы банковской статистики, а также подходы к ее формированию. В третьем разделе рассматривается процесс построения и улучшения модели вероятности дефолта российского банка, тестируется качество построенной модели, интерпретируются результаты. В четвертом разделе проведено сравнение качества построен-

ной и альтернативных моделей вероятности дефолта банка. В Заключении приведены выводы по данной работе.

1. Предпосылки построения моделей вероятности дефолта для российских банков

1.1. Российская банковская система и проблема устойчивости банков

Коммерческие банки стали возрождаться в России с 1988 г. До 1995 г. было создано более 2500 коммерческих банков, а всего в российской банковской системе (РБС) выдано около 3500 лицензий на деятельность кредитных организаций (КО) (табл. 1).

Таблица 1. Количественная динамика формирования российской банковской системы

№	Характеристика	По состоянию на 1 декабря соответствующего года									
		1997	1998	2000	2002	2004	2006	2008	2010	2011	2012 ¹
1	Зарегистрировано КО	2552	2481	2124	1826	1516	1345	1228	1146	1117	1 102
1.1	в т.ч. банков	2526	2451	2084	1773	1464	1293	1172	1084	1055	1 036
1.2	в т.ч. банков со 100% иностранным участием	16	18	22	27	33	52	77	80	77	76
2	Банки, имеющие право на банковские операции	1654	1447	1274	1282	1249	1143	1058	965	925	905
2.1	в т.ч. генеральную лицензию	262	263	244	293	311	287	298	283	273	271
2.2	в т.ч. лицензию на вклады населения	1589	1372	1239	1202	1165	921	886	819	799	786
3	КО, у которых отозвана лицензия	852	1004	806	491	218	155	117	132	135	136
4	КО, где принято решение о ликвидации	408	468	869	1238	1569	1758	1900	1991	2023	2 043
4.1	в т.ч. в форме присоединения	319	326	340	357	367	389	402	433	448	454

Источник: [www.cbr.ru].

¹ По состоянию на 1 августа 2012 г.

Можно выделить периоды интенсивного (1996–2000 гг.) и более быстрого (2008–2010 гг.) роста количества отозванных лицензий. Это периоды, соответствующие кризисам 1998 г. и 2008 г. По приведенным в табл. 1 показателям можно судить как о глубине кризиса, так и о тенденции к упорядочению банковской деятельности за счет регуляторных решений.

Начиная с 2000 г. существенно снизилось число вновь регистрируемых банков. Снижение количества действующих банков за прошедшее десятилетие связано как с выходом из бизнеса мелких банков за счет повышения регуляторных требований к капиталу, так и с реорганизацией деятельности ряда средних и крупных кредитных организаций.

В результате последовательных преобразований, законодательных и регуляторных новаций за 20 лет деятельности РБС сформирована рыночная территориально распределенная банковская система. Период экстенсивного формирования двухуровневой РБС завершен по выходу из кризиса 1998 г.

Наметился явный рост РБС в мировой банковской системе. До кризиса 2008 г. 35 российских банков попали в мировую Top-1000 (в 2000 г. – 20). Два из них (Сбербанк РФ и ВТБ) входят в первую мировую сотню. Активы российских банков соизмеримы с активами стран БРИК (более 0,5% от активов Top-1000 крупнейших мировых банков у каждой из стран) за исключением Китая (5%). Но российские банки достаточно серьезно уступают крупнейшим банкам мира, которые на порядок больше по объемным показателям, чем российские лидеры, а активы всей РБС меньше активов любого банка из числа мировой Top-20. Достигнутый уровень развития РБС, несмотря на высокие темпы роста, не полностью соответствует потребностям российской экономики. Многие предприятия вынуждены развиваться за счет собственных средств или зарубежных заимствований.

Российская банковская система в своем современном развитии прошла два полных этапа, водоразделами для которых стали крупнейшие кризисы 1998 г. и 2008 г. Достигнут уровень развития, обеспечивающий удовлетворительное выполнение традиционных банковских посреднических функций по финансированию российской экономики и домашних хозяйств.

Если *этап становления* (1989–1999 гг.) характеризовался спонтанным развитием РБС, созданием избыточного количества (порядка

3000) банков, то в 2000 г. число действующих КО составляло менее 1300, а к середине 2012 г. банковские лицензии имели немногим более 900 организаций.

Этап динамичного развития (2000–2008 гг.) отличался ускоренным наращиванием количественных и частично качественных показателей РБС. Достаточно успешно были преодолены многие структурные проблемы, выявленные кризисом 1998 г. Среднегодовые реальные темпы роста составляли более 20% по активам и капиталу, более 30% по кредитам, что существенно выше темпов роста ВВП (рис. 1). К концу 2008 г. активы составляли 67%, капитал – 9%, а кредиты – 48% от ВВП (рост в 2–4 раза по сравнению с показателями 2000 г.), что все еще недостаточно, но уже соизмеримо с показателями стран Центральной и Восточной Европы.

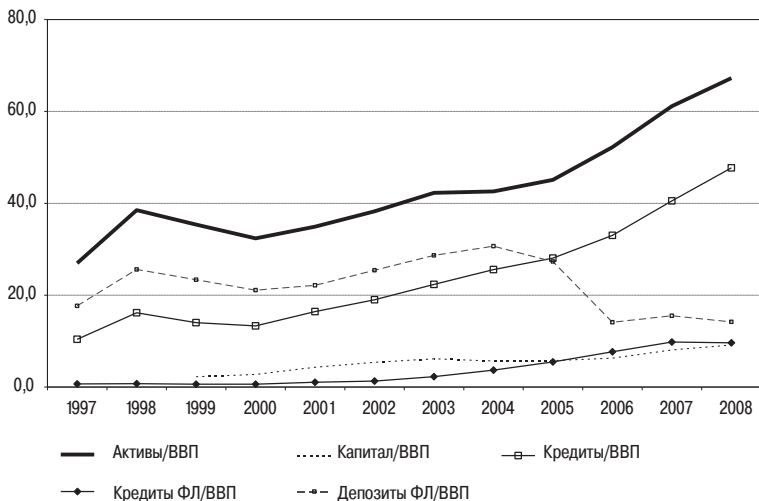


Рис. 1. Характеристики банковской системы России в сопоставлении с ВВП

Несбалансированный рост объемов кредитования сформировал признаки потенциального «перегрева» в предоставлении кредитных услуг, вызванные значительными временными разрывами и диспропорциями между активами и ресурсной базой, а также структурой привлеченных средств. Как следствие – рост просроченной и проблемной задолженности банков в условиях кризиса 2008 г.

Каждый кризис таит в себе потери и новые возможности. Он как минимум заставляет осмыслить текущее состояние, оценить проблемы, наметить пути их решения в системной интерпретации, во взгляде на задачи экономики в целом и роль РБС в ней. Представляется, что как в России, так и за рубежом дальнейшее развитие будет ориентировано на обеспечение устойчивости при достаточном уровне прибыльности.

Этап устойчивого развития (начиная с 2010 г. — момента начала выхода из кризиса 2008–2009 гг.) предусматривает смещение акцентов с вопросов наращивания объемов на управление балансом и динамикой развития. Среди основных проблем этого этапа можно выделить обеспечение пропорционального развития РБС в интересах кредитования реальной экономики и населения, решение ресурсной проблемы, укрепление капитальной базы РБС, а также решение проблемы эффективного регулирования финансовых институтов (включая банки и банковские холдинги) на национальном и наднациональном уровнях.

Кризис 2008 г. показал, что причины его возникновения за рубежом и распространения в развивающихся странах, включая Россию, лежат в регуляторной сфере. Вопрос о координации регулирования актуален как внутри страны, так и между государствами. Базель II не дает исчерпывающего ответа на эти вопросы, хотя в странах Западной Европы, где его механизмы используются более широко, кризисные явления в финансовой сфере не столь остры. В рамках институционального развития РБС важным является развитие *скоординированного управления и надзора*.

Важнейшая проблема заключается во внедрении современных систем риск-менеджмента. Это не только методологическая и организационная, но и достаточно сложная технологическая задача. В частности, отзыв в последние годы порядка 40 лицензий ежегодно способствует «расчистке» банковского сектора Банком России. Одной из основных задач по реформированию РБС является формирование системы мониторинга за деятельностью отдельных банков и РБС в целом.

Несмотря на устойчивую тенденцию к снижению количества банков, по состоянию на начало 2012 г. в России действует 978 кредитных организаций, 56 из них небанковские. За 15-летний период

с начала 1998 г. по июнь 2012 г. лицензии на осуществление банковской деятельности были отозваны у 926 кредитных организаций.

После финансового кризиса 1998 г. и 1999 г., когда банковские лицензии потеряли 14% и 9% всех национальных КО соответственно (рис. 2), ежегодная доля теряющих лицензию КО в основном не превышала 5%-го уровня за исключением 2005–2006 гг. (в результате ужесточения надзора за отмыванием средств) и 2009 г. (в результате недавнего мирового финансового кризиса). Очевидно, что каждый отзыв лицензий ставит под угрозу клиентов и партнеров банка, не способствует стабильности банковской системы в целом.

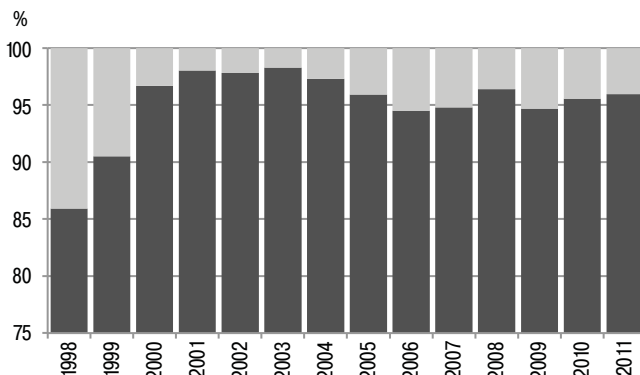


Рис. 2. Соотношение успешно функционирующих КО и КО, у которых отозвали лицензию, относительно числа КО, имеющих право на осуществление банковских операций на конец 2011 г.

Банк России как главный банковский регулятор заинтересован в обнаружении КО с неустойчивым финансовым положением заблаговременно: это позволит регулятору применять доступные ему инструменты для стабилизации ситуации и предотвращения потенциального отзыва лицензии в ближайшем будущем, минимизации потерь от возможного дефолта, по возможности, не только за счет регулярных выездных проверок.

Для этого необходима разработка эффективной системы, позволяющей регулятору дистанционно выявлять наиболее уязвимые кредитные организации для их более тщательного мониторинга. Качественная модель вероятности дефолта КО предоставит такие возможности.

Кроме того, модель вероятности дефолта может использоваться для оценки рисков, которые несет в себе банковская система (например, через оценку средней вероятности дефолта по 100 крупнейшим по величине активов банкам). На сегодняшний день требование к достаточности капитала — один из важнейших нормативов ЦБ РФ, обеспечивающий стабильность банковского сектора. В соответствии с Базель II, действующим сегодня глобальным банковским соглашением, минимальные требования к достаточности капитала — одна из основ стабильности банковской системы. Изменяя эти требования, регулятор изменяет необходимый, по его мнению, «запас прочности банка», созданный для покрытия различных рисков банковской деятельности.

Очевидно, что требования достаточности капитала должны быть жестче, когда риски банковской деятельности выше. На практике это приводит к тому, что в период рецессии, когда ранее накопленные риски материализуются, растут требования к достаточности капитала банков, углубляя колебания делового цикла. Эта проблема известна как проблема процикличности требований к капиталу. В действительности повышение требований к капиталу необходимо проводить именно в период накопления рисков (в конце подъема), а не их материализации (в начале рецессии), несмотря на политические сложности реализации этих действий. Не исключено, что динамика вероятностей дефолта банков подскажет ЦБ, в какой именно момент необходимо повышать требования к капиталу. Иными словами, адекватная модель вероятности дефолта поможет бороться с проблемой процикличности.

Таким образом, модель вероятности дефолта может быть востребована тремя основными агентами: Банком России, коммерческими банками и их кредиторами. При помощи модели регулятор сможет выявлять наиболее уязвимые банки (группу риска) и своевременно принимать меры для их финансового оздоровления. Наблюдение за динамикой вероятности дефолта, своей и контрагентов, позволит коммерческим банкам получить независимую оценку их устойчивости и перспектив развития, а также количественно оценивать влияние своих действий на риски, связанные с банковской деятельностью. Модель может оказаться полезной и для кредиторов банка, помогая им лучше осознавать риски кредитования конкретного банка.

1.2. Моделирование вероятности дефолта: обзор литературы

Подходы к построению моделей раннего предупреждения для банковского сектора и факторы, которые определяют успешность операционной деятельности банков в развитых странах, рассмотрены в ряде работ, обобщенных в (Bluhm et al., 2010). Далее мы сосредоточимся на рассмотрении преимущественно тех работ, в которых исследуется опыт России и развивающихся стран. По нашему мнению, такой подход позволит лучше учесть национальную специфику банковской деятельности в России, а также особенности уже проведенных исследований схожей тематики. Мы начнем с работ зарубежных авторов, но и в России проведен ряд интересных исследований. Среди них можно выделить исследования (Карминский и др., 2005; Пересецкий, 2010; Peresetsky et al., 2011), опыт которых в значительной степени используется в данной работе.

Финансовые переменные в наибольшей степени характеризуют устойчивость позиции банка, в том числе такие переменные, как размер активов банка и его капитализация (Peresetsky et al., 2011). Первый фактор чаще всего измеряется показателем «натуральный логарифм суммарных активов банка», второй — отношением капитала банка к суммарным активам. Размер банка является достаточно важной характеристикой, так как:

- показатель размера банка является значимым практически во всех моделях вероятности дефолта (Peresetsky et al., 2011);
- доля долгосрочных кредитов на срок более 3 лет в активах банка напрямую зависит от размера банка, согласно выводам (Chernykh, Theodossiou, 2011);
- крупным частным банкам присущ более высокий риск несостоятельности (insolvency risk), что нашло подтверждение в работе (Fungacova, Solanko, 2009);
- дефолты крупных банков приводят к значительному стрессу в экономике, чего регулятор стремится избежать, причем иногда он не в состоянии разобраться со сделками, которые проводят крупные банки (Claeys, Schoors, 2007).

Достаточность капитала (в России — норматив достаточности капитала) характеризует уровень средств банка, которыми он рискует при осуществлении операционной деятельности. Наша гипотеза состоит в том, что этот коэффициент должен входить в модель нели-

нейным образом. У банков с низким значением показателя достаточности капитала может оказаться недостаточно собственных средств для покрытия принятых рисков. У банков с высоким значением данного показателя бизнес-модель операционной деятельности построена неэффективно и не использует имеющихся возможностей долгового финансирования. Согласно статье (Tabak et al., 2011), в которой рассматривается банковская система Бразилии, подобная неэффективность приводит к тому, что у банка увеличивается доля неработающих активов. Принимая во внимание предыдущие два вывода, можно сказать, что существует U-образная зависимость между нормативом достаточности капитала и вероятностью дефолта банка.

В работе (Lanine, Venet, 2006) рассматриваются эмпирические модели раннего предупреждения для российского банковского сектора. Авторы отмечают, что большее отношение капитала к активам способствует снижению вероятности дефолта кредитной организации, в то время как размер банка, в отличие от указанных ранее работ, не влияет на вероятность его дефолта.

Следующим фактором, который потенциально влияет на вероятность дефолта, является ликвидность банка. Интересно, что по результатам проведенного в работе (Lanine, Venet, 2006) анализа отношение ликвидных средств банка увеличивает вероятность его дефолта. При этом отмечается, что данный факт имеет теоретическое и эмпирическое подтверждение. Ликвидные средства необходимы банку, чтобы своевременно и без значительных финансовых потерь удовлетворять требования вкладчиков и кредиторов. С другой стороны, значительное отвлечение ресурсов в ликвидные активы не способствует росту отдачи от инвестирования.

Для анализа влияния ликвидности на вероятность дефолта банка необходимо определить потенциальные уязвимости, связанные с ликвидностью банка. В этой связи представляет интерес вопрос, приводит ли наличие достаточно большого количества негосударственных ценных бумаг в активах банка к увеличению вероятности его дефолта.

Также в литературе уделяется внимание учету профиля риска банка при помощи Z-индекса. Он показывает, на сколько своих стандартных отклонений должна упасть прибыльность банка для того, чтобы истощить его капитал. В исследовании (Fungacova, Solanko, 2009) Z-индекс рассматривается в качестве непосредственной меры

риска, который принимает на себя банк, определяющей вероятность его несостоятельности. Расчет Z-индекса будет произведен в соответствии с подходом, реализованным в этом исследовании.

По общему мнению исследователей, включение в исходную модель макроэкономических переменных позволяет улучшить прогнозное качество исходной модели (Peresetsky et al., 2011). В данной работе в качестве макропеременных будут рассмотрены темпы роста ВВП и уровень инфляции, поскольку данные показатели являются опережающими индикаторами банковского кризиса. В статье (Vock, Demyanets, 2012) авторы строят панельную регрессию для определения ключевых факторов, влияющих на уровень просроченной задолженности в развивающихся странах. Согласно панельной регрессии, наблюдается отрицательная взаимосвязь между объясняемой переменной и темпами роста ВВП. В другой статье (Mannasoo, Mayes, 2009) строится модель вероятности дефолта банков в странах Восточной Европы. Авторы отмечают, что темпы роста ВВП являются одним из ключевых опережающих макроэкономических индикаторов, который должен использоваться в моделях оценки вероятности дефолта банков. В работе (Карминский и др., 2005) продемонстрировано, что параметры, отражающие общее улучшение в экономике, оказывают положительное влияние на состояние банка. Рост ВВП, рост промышленного производства, увеличение реальных доходов населения и улучшение внешнеторговой ситуации являются параметрами, которые оказывают потенциально положительное влияние на устойчивость банка в долгосрочной перспективе.

Следующая группа факторов, которые часто используются для улучшения моделей вероятности дефолта, – институциональные признаки банков.

Согласно ряду статей, тип собственности в банковской системе РФ является одной из ключевых характеристик банка. Например, в статье (Fungacova, Solanko, 2009) делается вывод о том, что у банков с иностранным участием в капитале более высокий риск несостоятельности, а у государственных – меньший по сравнению с остальными банками. К противоположному выводу пришли авторы статьи (Micco et al., 2007). Согласно проведенному исследованию, иностранные банки достигают лучших операционных результатов, чем национальные банки.

В статье (Clarke et al., 2005) названы три главных причины, по которым государственные банки могут проигрывать частным, а значит, и иметь повышенную вероятность дефолта: вмешательство политиков в дела банка; проблемы с менеджментом, остро стоит агентская проблема; нехватка конкурентной среды (госбанки часто защищены от чистой конкуренции, субсидируются государством). Идея неэффективности госбанков находит подтверждение в статье (Misso et al., 2007): в развивающихся странах государственные банки достигают скромных результатов операционной деятельности, нанимают избыточное количество персонала и несут большие административные издержки. С другой стороны, помощь от государства в случае возникновения финансовых трудностей и расширенный доступ к рынку МБК приводят к снижению вероятности дефолта государственного банка.

К сожалению, при помощи регрессионного анализа невозможно оценить влияние типа собственности банка на вероятность дефолта кредитной организации. Дело в том, что ни один банк со 100%-м участием нерезидентов в нашей выборке не допустил дефолта за 1998–2011 гг. Некоторое снижение числа банков со 100%-м участием нерезидентов в капитале произошло в результате мирового финансового кризиса 2008–2009 гг. (рис. 3), однако отзывы лицензий были связаны с реорганизацией банков.

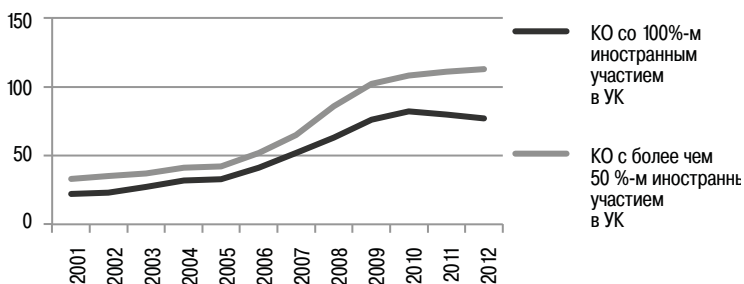


Рис. 3. Динамика численности КО со 100%-м и 50%-м участием нерезидентов в капитале, 2001–2012 гг., на начало года

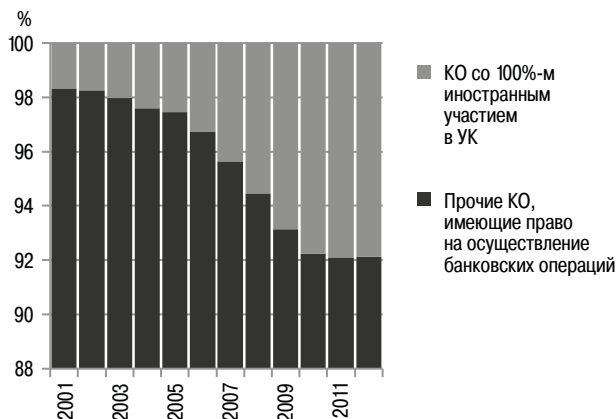


Рис. 4. Соотношение КО со 100%-м участием нерезидентов в УК и прочих КО, 2001–2012 гг., на начало года

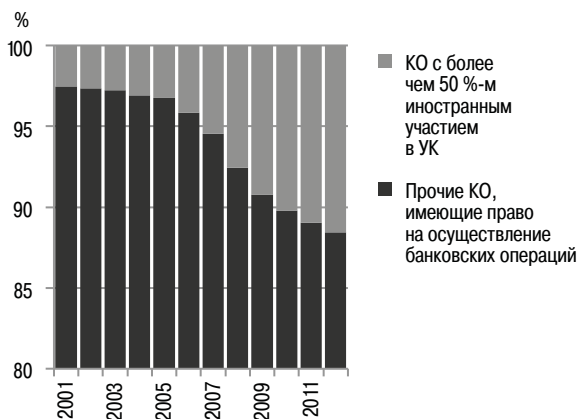


Рис. 5. Соотношение КО с более чем 50%-м участием нерезидентов в УК и прочих КО, 2001–2012 гг., на начало года

Тем не менее в условиях растущей экономической нестабильности в мировых масштабах 100%-е участие нерезидентов в капитале банка не гарантирует ему безоблачного финансового состояния (и соответственно, низкую вероятность дефолта). В этой связи банки с иностранным участием требуют не менее пристального внимания со стороны регулятора.

Возможным выходом из сложившейся ситуации является учет в качестве банков с иностранным участием КО с более чем 50%-м участием нерезидентов в капитале, которых несколько больше в российской экономике (рис. 4, 5). В случае, если среди них окажутся потерпевшие дефолт КО, регрессионный анализ теоретически применим. Однако на сегодняшний день не существует открытого источника, содержащего достоверную и полноценную информацию такого рода.

Также в российской банковской практике отсутствуют случаи дефолта банков с высоким участием государства в капитале (более 50%). Перечень государственных банков за рассматриваемый период времени был позаимствован из исследования (Vernikov, 2011). Государство склонно финансово поддерживать банки, в капитале которых оно участвует. Это поведение соответствует поведению разумного собственника. Количественно учесть данное влияние не представляется возможным.

В работе (Карминский и др., 2005) была осуществлена кластеризация банков, т.е. выделялись группы банков, для которых строились модели оценки вероятности дефолта. Это оказало лишь скромный положительный эффект на качество построенных моделей.

На деятельности банка может отражаться факт его участия в системе страхования вкладов (ССВ). Если банк является членом ССВ, есть вероятность того, что он будет принимать на баланс больше риска (Demirguc-Kunt, Huizinga, 2004). При наличии ССВ банкам легче получать денежные средства от вкладчиков, впоследствии не интересующихся финансовым состоянием банка. Они ориентируются лишь на размер процента, под который можно разместить деньги. В результате банки имеют дополнительную мотивацию к проведению рискованных операций. Дамми-переменная на участие в ССВ может служить прокси-показателем морального риска (*moral hazard*).

Следующий институциональный фактор – расположение главного офиса банка. В статье (Claeys, Schoors, 2007) подчеркивается, что Банк России менее склонен отзываться лицензии у региональных банков в регионах с малым присутствием кредитных организаций (чтобы не ослаблять и без того умеренную конкуренцию на региональных рынках).

Уровень конкуренции в банковском секторе существенно различается по регионам (Anzoategui et al., 2012). На локальных конкурент-

ных рынках для банков выше вероятность снижения доходов, что может привести в дефолту банка. Рыночная структура банковского сектора и уровень конкуренции в нем чаще всего отражаются при помощи двух индексов: Лернера и Херфиндала. Оба показывают степень монополизации, однако индекс Лернера показывает уровень монопольной власти отдельно взятой фирмы, в то время как индекс Херфиндала – уровень монополизации отрасли в целом. В статье (Fungasova, Weill, 2009) установлена обратная связь между индексом Лернера и вероятностью дефолта банка. При получении данного вывода авторы воспользовались логистической регрессией с панельной структурой данных.

Подробный обзор различных типов моделей вероятности дефолта выходит за рамки данной работы. Полноценное освещение данного вопроса произведено в статье (Тотьмянина, 2011).

2. Формирование базы данных для эмпирического моделирования вероятности дефолта банка

2.1. Понятие и классификация видов дефолта

Основной фокус данной работы заключался в построении моделей отзыва лицензии вследствие неплатежеспособности банка и отсутствия необходимых собственных средств банка на покрытие рисков, связанных с проведением активных операций. Необходимость четкого определения дефолта банка основывается на том, что в исходную выборку попадают банки, у которых лицензия была отозвана из-за махинаций (отмывание денежных средств, пособничество терроризму), а также по собственной инициативе. Класс моделей, описывающий махинации, выходит за рамки обсуждения данной работы. Особенности моделирования отзыва лицензий российских банков по прочим причинам² приведены в статье (Пересецкий, 2010).

² Перечень причин, по которым отзываются лицензии у банка, приведен в Федеральном законе «О банках и банковской деятельности» от 02.12.1990 N 395-1, ст. 20.

Поэтому в данной работе вводится следующее определение дефолта банка: банк считается обанкротившимся тогда и только тогда, когда:

- достаточность собственного капитала становится ниже 2%;
- размер собственных средств ниже минимального значения уставного капитала на дату регистрации банка;
- банк не исполняет в срок, установленный Федеральным законом «О неисполнении (банкротстве) кредитных организаций», требования Банка России о приведении в соответствие величины уставного капитала и размера собственных средств капитала;
- банк не способен удовлетворить требования кредиторов по денежным обязательствам и (или) исполнить обязанность по уплате обязательных платежей;
- банк был подвержен санации.

Данные о фактах и причинах отзывов лицензий российских банков собирались с использованием приказов об отзыве (аннулировании) лицензий на осуществление банковских операций, изданных Центральным банком РФ. Информация о проведенных санациях банков доступна на официальном сайте АСВ (АРКО), а также в других открытых источниках. В случае, если между санацией банка и отзывом его лицензии, признанным дефолтом, прошло менее двух лет, данные события рассматривались как единый дефолт, произошедший в более ранний из двух возможных моментов времени.

2.2. Источники национальной банковской статистики

При подготовке данного исследования нами были проанализированы три основных источника финансовой отчетности банков. Базы данных по финансовой отчетности строятся на основе отчетов о балансовых данных и о прибылях и убытках (101 и 102 формы в соответствии с инструкцией Банка России). Более детальные характеристики каждой из баз данных приведены в табл. 2.

С учетом предварительного анализа и задач исследования, информация о финансовых показателях была взята из ИАС «Банки и финансы» (продукт Информационного агентства «Мобиле»). Основным фактором нашего выбора является то, что временной ряд здесь шире по сравнению с базой Информационного агентства Интерфакс. Рассматривать более ранний период времени (до 1998 г.) не имеет смыс-

ла, так как в 1998 г. произошло существенное изменение в бухгалтерском учете российских банков, а также до 1998 г. в РБС имелось большое число «мелких», «спящих» банков.

Таблица 2. Характеристики баз данных по финансовой отчетности банков на территории РФ

Название	Степень покрытия российского банковского сектора	Период покрытия	Частота представления данных	Наличие пропусков
ИАС «Банки и финансы»	Высокая	1998–2012 гг.	Месяц	Много
BankScore	Низкая	1996–2012 гг.	Квартал	Мало
Интерфакс	Высокая	2000–2012 гг.	Квартал	Много

В нашем распоряжении находилась именно база данных «Банки и финансы», за что авторы признательны проф. А.Е. Петрову.

2.3. Описание сформированной базы финансовой статистики

Для каждого банка рассматривались поквартальные данные за период 1998–2011 гг. Использование помесечных данных представляется нецелесообразным в силу их «замусоренности».

Возможны два состояния банка: «0» — банк продолжает операционную деятельность (выжил); «1» — банк находится в состоянии дефолта. Дефолт регистрировался в том случае, если среди причин отзыва лицензии имелась хотя бы одна из перечисленных в определении, приведенном в п. 2.1.

За рассматриваемый 14-летний период (1998–2011 гг.) было зафиксировано 910 фактов отзыва лицензии, а также 37 случаев, когда банки были подвержены санации. Динамика отзывов лицензий у КО в сравнении с теми отзывами, которые признаны дефолтными, приведена на рис. 6; а отзывов, связанных с санацией банка, — на рис. 7. По большинству из этих банков в использовавшейся базе данных была представлена финансовая статистика (рис. 8).

Динамика отзывов лицензий у банков в целом соответствует динамике отзывов, признанных дефолтами в соответствии с нашей классификацией, с тем исключением, что в 2005–2007 гг. отсутствует увеличение числа истинных дефолтов при увеличении числа отзывов:

как уже было сказано ранее, отзывы были связаны с ужесточением надзора в целом.

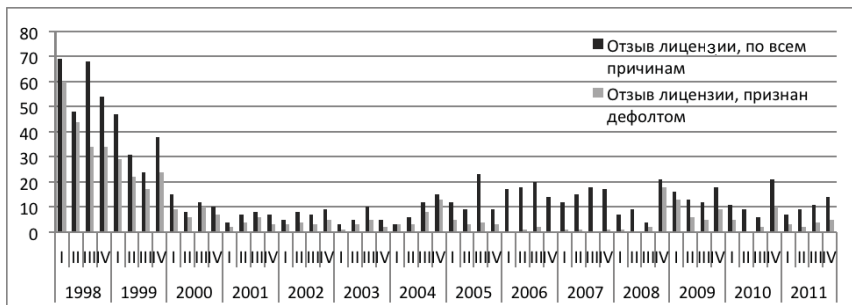


Рис. 6. Сравнение исторической динамики отзывов лицензии у кредитных организаций и отзывов, признанных дефолтами, поквартально, 1998–2011 гг.



Рис. 7. Дефолты, связанные с санацией банков. Распределение произошедших санаций банков и их покрытие использованной базой данных «Банки и финансы», поквартально, 1998–2011 гг.

Для проведения исследования было решено использовать квази-панельную структуру данных, которую легко трансформировать в панельную. В табл. 3 представлена структура базы данных.

Можно отметить высокую степень несбалансированности данных в том смысле, что наблюдений для обанкротившихся банков существенно меньше, чем для успешно функционирующих.

Таблица 3. Структура базы данных. Типичное наблюдение

Номер лицензии банка_период	Состояние банка	Набор финансовых показателей, тыс. руб.				
		BP	CP	...	KE	LA
507_1/4/2005 ³	0	219	115	...	83513	31830

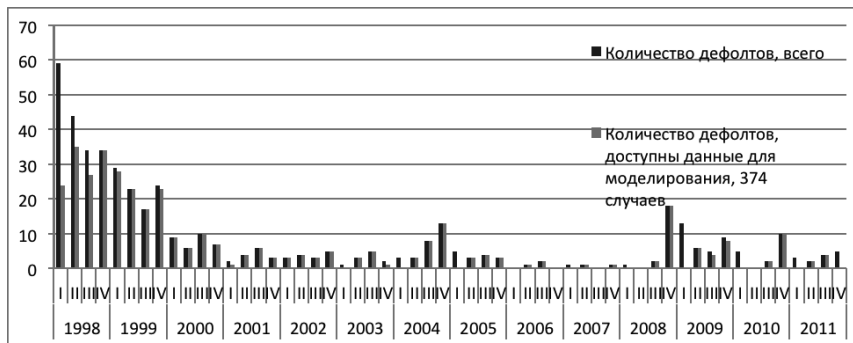


Рис. 8. Сравнение исторического числа дефолтов и случаев дефолта кредитных организаций, по которым имелись финансовые показатели на момент сбора данных (конец 2011 г.), поквартально, 1998–2011 гг.

В ИАС «Банки и финансы» информация по всем переменным изначально приведена помесечно. Перечень используемых переменных и их описательные статистики содержатся в Приложении 1.

Значения последних семи переменных (начиная с ODB_q и заканчивая RUB2_q) брались из отчетов о прибыли и убытках (форма 102) соответствующих банков, в которых учет в течение года происходит накопительным образом, поэтому в базе данных для исследования использовался прирост значения этих переменных за соответствующий квартал, а не абсолютные значения.

Все остальные переменные (BP – NORM_H3) сформированы на основе бухгалтерских балансов по российской отчетности. Поэтому в базе данных для исследования использовались значения этих показателей на конец квартала.

³ 507_1/4/2005 – 1 апреля 2005 г., конец I кв. 2005 г.; наблюдение по банку с номером лицензии 507.

Для целей построения моделей вероятности дефолта банка исходная выборка была разбита на две части: первую, по которой строятся модели (наблюдения с 1998 по 2009 г.) и вторую (наблюдения с 2010 по 2011 г.), по которой будет оцениваться прогнозная сила построенных моделей.

2.4. Описание модели бинарного выбора: *logit-модель*

Формирование данных для эмпирического исследования проводилось с учетом способа их последующего оценивания при моделировании вероятности дефолта. В данной работе для расчета вероятности дефолта была использована *logit-модель* (Магнус и др., 2007).

Вероятность дефолта банка (т.е. события $default = 1$) с использованием *logit-модели* рассчитывается следующим образом:

$$P(default = 1) = \Lambda(x \cdot \beta), \text{ при } 0 \leq \Lambda() \leq 1,$$

$$\text{при этом } \begin{cases} P(default = 1) \rightarrow 1 \text{ при } x \cdot \beta \rightarrow \infty \\ P(default = 1) \rightarrow 0 \text{ при } x \cdot \beta \rightarrow -\infty \end{cases}$$

При таком подходе вероятность дефолта не может выходить за рамки отрезка $[0; 1]$, а также предполагается нелинейность зависимости вероятности дефолта от используемых объясняющих факторов.

В случае логистической модели функциональная часть имеет вид:

$$\Lambda(x \cdot \beta) = \frac{\exp(x \cdot \beta)}{1 + \exp(x \cdot \beta)}.$$

Для оценки моделей применяется метод максимального правдоподобия. В качестве ML-оценки вектора коэффициентов β принимается такой вектор $\hat{\beta}$, для которого вероятность наблюдения имеющихся значений y (состояний банка) максимальна:

$$L(y) = \prod_{i=1}^v [\Lambda(x_i \beta)] \cdot \prod_{j=1}^w [1 - \Lambda(x_j \beta)],$$

при том, что выборка состоит из v наблюдений $default = 1$ и w наблюдений $default = 0$.

Тогда искомым вектор $\hat{\beta}$ находится из следующей задачи:

$$\ln[L(y)] \rightarrow \max_{\beta}.$$

Как уже было сказано, можно наблюдать только два состояния банка: дефолт или продолжение операционной деятельности. Для причисления банка к одному из классов вводится ненаблюдаемая переменная $default^*$. Эта переменная считается непрерывной и определяется моделью:

$$default^* = \Lambda(x' \cdot \beta).$$

Предсказание состояния банка определяется моделью по следующему правилу:

$$\begin{cases} default = 1 \text{ при } c \leq default^* \leq 1 \\ default = 0 \text{ при } 0 \leq default^* \leq c \end{cases}, \text{ где } c \subseteq [0;1],$$

где критерий разделения классов «с» задается пользователем модели.

3. Построение модели бинарного выбора для вероятности дефолта банка

3.1. Первичный отбор и очистка объясняющих переменных

Смысловая и статистическая очистка. Информационно-аналитическая система «Банки и финансы» включает порядка 170 финансовых показателей, определяемых на основе российской отчетности коммерческих банков. В то же время не все показатели на протяжении анализируемого временного интервала представлены в базе на презентативном уровне. В этой связи оставлены только показатели, перечисленные в Приложении 1.

Кроме того, для построения модели лучше использовать не сами финансовые индикаторы, а некоторые производные от них. В этой связи первоначально был сформирован набор возможных объясняющих переменных для построения модели вероятности дефолта банка (относительных финансовых переменных, оцененных на основе абсолютных значений, а также логарифма чистых активов, используемого в качестве показателя размера банка). Выбор объясняющих

переменных осуществлен исходя из возможности базы, накопленного в других исследованиях опыта на основе анализа, представленного в п. 1.2, а также анализа объясняющих возможностей соответствующих переменных статистическими методами.

Очистка данных была проведена в несколько этапов. На первом этапе отсеивались явные ошибки (ввода или измерения): были удалены наблюдения с отрицательными значениями по переменным *чистые активы CA*, *кредиты экономики KE*, *капитал SK*, *норматив текущей ликвидности NORM_H3*. Затем по каждой из относительных переменных были удалены наблюдения, содержащие статистические выбросы, причем только для банков, продолжающих операционную деятельность. Наблюдения удалялись только для функционирующих банков, поскольку в случае дефолта финансовые показатели банков могут иметь сомнительные значения в силу слабого финансового состояния и недостоверной финансовой отчетности. При этом в данной работе любая операция по очистке влечет за собой аналогичные действия по соответствующим лаговым переменным. Например, удаление отрицательных наблюдений для переменной *CA* предполагает удаление отрицательных значений по переменным *CA* с лагами от 1 до 8 кварталов, т.е. для переменных, далее обозначаемых как *CA_lag1*, *CA_lag2*, *CA_lag3*, ... *CA_lag8*.

Тест на разделительную способность (ANOVA-тест). Для предварительного формирования и очистки объясняющих переменных применительно к модели вероятности дефолта банков был проведен тест на разделяющую способность переменных (ANOVA-тест). Чтобы воспользоваться данным тестом, были сформированы средние значения финансовых показателей по группам функционирующих и объявивших дефолт банков. Основным смысл ANOVA-теста⁴ заключается в том, чтобы выделить те финансовые переменные, для которых средние значения по двум группам банков статистически отличаются. Затем среди переменных с высокой разделяющей способностью⁵ были выявлены и отсеяны те, по которым имеется недостаточное или не-

⁴ Для сохранения переменной нулевая гипотеза (о равенстве средних значений между двумя группами) проводимого теста должна быть отвергнута и, соответственно, принята альтернативная гипотеза (о различии средних значений).

⁵ Гипотеза о равенстве средних значений для групп выживших и допустивших дефолт банков была отвергнута на уровне значимости 10%.

равномерно распределенное количество наблюдений на протяжении 1998–2011 гг. Результаты данного теста приведены в табл. 4.

Исследование парных корреляций отобранных переменных. При рассмотрении парных корреляций отобранных переменных в парах переменных, имеющих потенциально близкий экономический смысл, была выбрана одна из них. При этом не исключалось, что при необходимости в эконометрической модели будет опробована и статистически связанная переменная, если первая окажется незначимой.

В работе в качестве максимально допустимого уровня связанности между объясняющими переменными был принят уровень корреляции, равный 0,3. Для сохранения сильно коррелированных переменных, не имеющих ярко выраженной экономической связи, были предприняты попытки перехода к логарифмическому масштабу (использование не самой относительной переменной, а ее логарифма). Матрица корреляций предварительно отобранных объясняющих переменных представлена в Приложении 2.

Таблица 4. Результаты теста автоматического выбора переменных

№	Наименование переменной	Обозначение переменной	Формула для вычисления	P-value в ANOVA-тесте
1	Отношение собственных средств к чистым активам	sk_ca	sk/ca	0,031
2	Отношение балансовой прибыли к чистым активам	bp_ca	bp/ca	0,042
3	Отношение ликвидных активов к чистым активам	la_ca	la/ca	0,007
4	Отношение объема негосударственных ценных бумаг к чистым активам	ncb_ca	ncb/ca	0,004
5	Отношение просроченной задолженности по ссудам к кредитам экономике	pzs_ke	pzs / ke	0,008
6	Отношение оборотов по корреспондентским счетам к чистым активам	oks_ca	oks / ca	0,072
7	Отношение ликвидных активов к обязательствам до востребования	la_ov	la / ov	0,109

№	Наименование переменной	Обозначение переменной	Формула для вычисления	P-value в ANOVA-тесте
8	Логарифм чистых активов	lnca	$\ln(ca)$	0,079
9	Отношение операционных доходов банка к чистой прибыли	odb_cp	odb / cp	0,165
10	Отношение чистой прибыли к чистым активам	cp_ca	cp / ca	0,078
11	Отношение ликвидных активов к долговым обязательствам государственным и органов местного самоуправления	la_gdo	la / gdo	0,123
12	Отношение ликвидных активов к обязательствам на срок свыше 1 года	la_solong	la / so_long	0,243
13	Отношение долговых обязательств государственных и органов местного самоуправления к чистым активам	gdo_ca	gdo / ca	0,324
14	Отношение работающих активов к чистым активам	ra_ca	ra / ca	0,168
15*	Отношение депозитов физических лиц до 30 дней к депозитам физических лиц	vdf130_dfl	$vdf1_30 / dfl$	0,069
16	Отношение расходов на содержание аппарата к операционным расходам банка	rsa_orb	rsa / orb	0,654
17*	Отношение просроченной задолженности по ссудам к обязательным резервам в ЦБ РФ	pzs_orcb	$pzs / orcb$	0,098
18	Отношение резервов к чистым активам	res_ca	res / ca	0,023

Примечание. Заштрихованы номера переменных, обладающие высокой разделяющей способностью. Переменные, по которым имеется достаточное разделение, но недостаточное или неравномерно распределенное во времени количество наблюдений, помечены звездочкой.

В результате мы отказались от использования переменных *отношение ликвидных активов к чистым активам* la_ca (корреляционно тесно связана с br_ca), *отношение чистой прибыли к чистым активам* sr_ca (тесно связана с br_ca), *отношение просроченной задолженности по ссудам к резервам под проблемные активы* prs_res (тесно связана с prs_ke), *отношение просроченной задолженности по ссудам к чистым активам* prs_ca (тесно связана с pcb_ca). В работе также использовался логарифм переменной *отношение оборотов по корреспондентским счетам к чистым активам* oks_ca.

Матрица парных корреляций отобранных объясняющих переменных приведена в Приложении 1.

Экономический смысл отобранных финансовых переменных. Теперь разобьем оставшиеся отобранные (перспективные) объясняющие переменные по группам, каждая из которых характеризует состояние банка по соответствующему блоку в классификации CAMELS. Параллельно мы опишем наши ожидания относительно влияния данных переменных на вероятность дефолта банков.

1. Группа переменных, связанных с *капиталом* (C, capital). После предварительного отбора эта группа представлена переменной *отношение капитала к активам* sk_ca.

С одной стороны, высокая доля собственного капитала в чистых активах банка указывает на наличие запаса прочности у банка на случай возникновения финансовых потрясений, а также, косвенно, — на готовность собственника вкладывать средства в бизнес. С другой стороны, слишком большая доля собственного капитала в чистых активах свидетельствует об отсутствии достаточно эффективной посреднической деятельности, в том числе об отсутствии достаточного долгового финансирования у кредитной организации. Это, в частности, может быть связано с неуверенностью кредиторов в жизнеспособности банка в долгосрочной перспективе. Кроме того, высокое значение данного показателя может свидетельствовать о том, что менеджмент банка значительную часть чистой прибыли вынужден вносить в создание резервов, которые составляют часть капитала, что может отражать уязвимое финансовое положение банка.

Естественно, малое значение этого показателя увеличивает вероятность дефолта банка, что косвенно отражено в нормативных документах как российских (Инструкция № 110 (Инструкция, 2004)), так и международных (Базельское соглашение (Basel II, 2004)).

Таким образом, данный показатель характеризует риск потери (недостаточности) капитала, который принимает на себя банк. Мы ожидаем положительное влияние роста отношения sk_ca на вероятность дефолта банка при слишком больших значениях этого отношения и падения данного показателя при слишком малых, т.е. U-образную зависимость.

2. Группа переменных, связанных с *активами* (A, assets). Данная группа представлена двумя объясняющими переменными, одна из которых также характеризует размер банка. Это *логарифм чистых активов* ln_ca .

Крупные банки могут выигрывать за счет лучшей диверсификации своего бизнеса (территориально, по отраслям, по конкретным заемщикам и др.). При этом нередко отмечается склонность крупных банков к риску, уверенность в том, что даже в случае возникновения финансовых трудностей, они все равно будут поддержаны государством из-за их системной значимости (известная проблема «слишком большой, чтобы обанкротиться» (too big to fail)).

Таким образом, влияние переменной ln_ca на вероятность выживания банка неоднозначно. По нашему мнению, влияние скорее U-образное, так как в новейшей российской истории немало случаев отзыва лицензий у крупнейших банков (Карминский и др., 2005), особенно в условиях кризиса 1998 г., когда из 20 крупнейших банков примерно половина потерпела крах. Соответствующие примеры имеются и в условиях рецессии 2008 г., когда в процессе кризиса и после него дефолтное состояние наблюдалось у 4–5 банков из Top-50.

К этой же группе относится переменная *отношение просроченной задолженности по ссудам к кредитам экономики* pzs_ke . Высокий уровень просроченной задолженности относительно величины выданных кредитов, отражает неблагоприятное финансовое положение КО, а главное – низкое качество оценки рисков, связанных с операциями кредитной организации. В будущем просроченная задолженность может быть списана на резервы или убытки.

Таким образом, мы ожидаем положительного влияния роста значения pzs_ke на вероятность дефолта КО. В определенной мере указанный показатель характеризует и качество менеджмента в кредитной организации.

3. Группа переменных, связанных с *менеджментом* (M, management). Помимо указанной выше переменной pzs_ke , к этой группе относит-

ся также переменная, характеризующая *отношение оборотов по корреспондентским счетам за период времени к чистым активам* (в логарифмическом масштабе) \ln_oks_ca . Эта переменная является важнейшим индикатором банковской активности и стратегического поведения менеджмента банка.

Слишком высокое значение отношения может сигнализировать как о некоторой панике в банке, так и о повышении его деловой активности, в том числе по обменным операциям. Слишком низкое значение отношения говорит о «застое в кровеносной системе банка», о потенциальных проблемах с проведением платежей или даже о свертывании банковского бизнеса. Мы ожидаем отрицательное влияние данного показателя на вероятность дефолта банка. Вопрос о выборе логарифмического масштаба переменной связан с избавлением от высокого уровня корреляции с другими объясняющими факторами.

4. Группа переменных, связанных с *прибылью* (E, earnings). Эта группа представлена переменной *отношение балансовой прибыли к чистым активам* br_ca .

Отрицательное отношение балансовой прибыли за квартал к размеру чистых активов свидетельствует о том, что КО является убыточной по итогам квартала. Прибыль — основной источник средств для развития, роста капитала и формирования устойчивого развития банка. Ни одна фирма не может существовать в долгосрочной перспективе, неся постоянные убытки. В то же время очень высокая доля прибыли относительно чистых активов вызывает подозрения, что банк принимает на себя очень большие риски, которые в определенный момент могут реализоваться и привести к дефолту кредитной организации.

Характеристики прибыльности банка имеют непосредственное отношение к вопросу о вероятности дефолта банка, ведь прибыль может быть использована на покрытие любых реализовавшихся рисков. Поэтому мы ожидаем, что падение отношения br_ca ниже определенного уровня, как и рост выше определенного значения, оказывает отрицательное влияние на вероятность выживания КО. При умеренных положительных значениях рост br_ca повышает вероятность выживания банка. Таким образом, есть все предпосылки ожидать U-образной зависимости (как и в случае sk_ca), хотя в данном случае возможно и не слишком выраженной.

5. Группа переменных, связанных с *ликвидностью* (L, liquidity) и *чувствительностью* (S, sensitivity). В рамках данного исследования мы использовали переменную *отношение объема негосударственных ценных бумаг к чистым активам* `ncb_sa`.

Операции с ценными бумагами, прежде всего с негосударственными, сопряжены с повышенным уровнем рыночного риска. При этом инвестирование в ценные бумаги, в том числе в негосударственные, является важным компонентом управления ликвидностью банка. По нашему мнению, банки, которые слишком активно вовлечены в торговлю негосударственными ЦБ и имеют высокое отношение `NCB_SA`, с большей вероятностью допустят дефолт в будущем.

Дополнительно мы планировали рассмотреть влияние показателя *норматив текущей ликвидности*, но имеющиеся данные не позволяют это сделать на всем горизонте исследования. Кроме того, следует отметить, что фактор чувствительности к рынку трудно оценить применительно к российским банкам в силу того, что немногие из них имеют торгуемые ценные бумаги.

3.2. Проблема несбалансированности данных

Logit-модель предполагает наличие определенной сбалансированности в выборке для адекватного построения. Ее качество снижается при наличии несбалансированной выборки, в которой один класс наблюдений (в нашем случае – продолжающие деятельность банки) существенно доминирует над другим классом (в нашем случае – над допустившими дефолт банками).

Иными словами, модель должна обучаться на наблюдениях обоих классов (на допустивших дефолт и успешно функционирующих банках). Данная ситуация опасна тем, что статистические критерии будут неадекватно отражать качество модели, а следовательно, пострадает и прогнозная сила модели вероятности дефолта. Несмотря на усилия по сохранению наблюдений типа «дефолт» при формировании базы, их доля составляет лишь порядка 1%.

Существует несколько способов балансировки данных. Основные подходы основаны на следующих методах (He et al., 2009):

- увеличение числа наблюдений меньшего класса;
- уменьшение числа наблюдений большого класса;
- формирование случайной выборки, балансирующей классы.

Для того чтобы минимизировать искажения в данных, далее использовался третий способ балансировки. При этом для увеличения доли наблюдений типа «дефолт» при моделировании применялся следующий алгоритм. Формировалась 1000 подвыборок, каждая содержала 5% наблюдений типа «банк продолжает операционную деятельность» от начальной выборки и все имеющиеся наблюдения типа «дефолт». Это позволяет повысить долю допустивших дефолт банков до 10–12% в каждой из подвыборок, а значит, каждая из них окажется пригодной для построения logit-модели. Каждая из рассматриваемых далее спецификаций модели будет тестироваться по одной из подвыборок, а затем будет произведено обобщение коэффициентов модели по всем подвыборкам.

3.3. Выбор величины лага по объясняющим переменным

Основной задачей конструируемой модели является прогнозирование банковских дефолтов. Поэтому для целей построения модели финансовые показатели брались с лагами. Финансовые трудности банка достаточно быстро находят отражение в его балансе, поэтому не имеет смысла рассматривать лаги, превышающие восемь кварталов. Этому же мнению придерживаются в литературе по данной тематике (Peresetsky et al., 2004).

Отдельного внимания заслуживает то, каким образом рассматривались лаги: для каждого момента времени финансовому показателю банка с лагом n соответствует данный финансовый показатель этого банка n периодами ранее (см. рис. 9 для случая лага в 4 квартала).

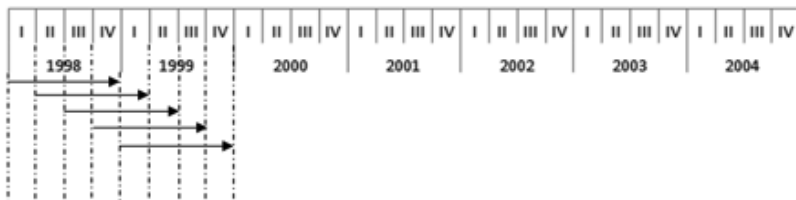


Рис. 9. Формирование лага в 4 квартала ($n = 4$) для финансовых показателей

Таким образом, необходимо выбрать величину лага для модели вероятности дефолта банка с обоснованной ранее спецификацией:

$$P(\text{default} = 1) = \Lambda(\text{sk_ca}; \ln_ca; \text{pzs_ke}; \ln_oks_ca; \\ \text{bp_ca}; \text{ncb_ca}).$$

Идея определения оптимального лага по финансовым переменным сводится к следующему:

1. Строятся логистические регрессии для каждой величины лага по финансовым переменным (от 1 до 8 кварталов, поквартально).
2. Анализируется динамика статистических характеристик спецификаций модели в зависимости от величины лага.
3. На основе анализа определяется величина лага по финансовым переменным.

При анализе статистических характеристик моделей здесь и далее рассматриваются следующие факторы:

- значимость коэффициентов при объясняющих переменных;
- уровень значимости моделей и значение соответствующего статистического показателя Pseudo R² (McFadden's);
- площадь под ROC-кривой, доминирование одних ROC-кривых, критерием для которых является мажорирование одной кривой над другими (мажорирование по Парето);
- статистические показатели (выборочные вероятности), описывающие адекватность модели эмпирическим данным:
 - доля верных предсказаний относительно всего количества наблюдений: *Pr* (верное предсказание);
 - доля верных предсказаний о дефолте банка относительно числа действительно обанкротившихся КО (sensitivity);
 - доля верных предсказаний о выживании банка относительно числа действительно выживших КО (specificity);
 - доля ложных предсказаний дефолта КО относительно общего числа предполагаемых дефолтов: *Pr* (ExcessWork).

Проведенный анализ показал, что с увеличением лага падает качество моделей и адекватность модели эмпирическим данным. Логично, что финансовые переменные непосредственно перед крахом (либо успешным продолжением операционной деятельности) позволят прогнозировать крах (продолжение деятельности) с большей точностью.

Тем не менее использовать лаг в один квартал недопустимо. Пользователь модели (в частности, регулятор) должен иметь возможность выявить банки из группы риска заранее, чтобы успеть предпринять меры, направленные на улучшение ситуации в банке. Как это показано на рис. 10, действительный промежуток между моментом дефолта и моментом измерения финансовых показателей может составить всего несколько дней: за это время невозможно эффективно вмешаться в деятельность ненадежного банка.

		Состояние банка		
		помесяч.	покварт.	
2001 год	I квартал	январь	0	0
		февраль	0	
		март	0	
	II квартал	апрель	1	1
		май		
		июнь		



Рис. 10. Демонстрация неадекватности использования лага в один квартал по объясняющим переменным: банк потерпел дефолт 3 апреля, однако при поквартальном измерении это будет учтено лишь в конце июня

С учетом приведенных соображений было принято решение об использовании в модели лага в два квартала как минимально допустимого. С учетом этих соображений скорректированная модель будет называться *базовой моделью* Basic, включать финансовые объясняющие переменные с лагом в два квартала и иметь следующую спецификацию:

$$P(\text{default} = 1) = \Lambda(sk_ca_{lag2}; \ln_ca_{lag2}; pzs_ke_{lag2}; \ln_oks_ca_{lag2}; bp_ca_{lag2}; ncb_ca_{lag2}).$$

На рис. 11 проведено сравнение ROC-кривых для моделей с различными лагами. Сводные статистические характеристики базовой модели Basic (как и последующих моделей) приведены в подразд. 3.6.

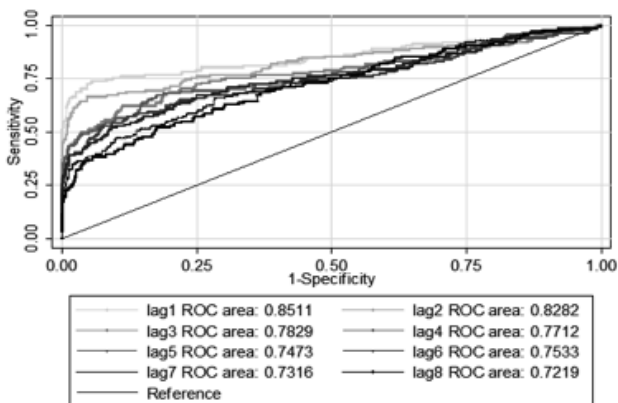


Рис. 11. Сравнение ROC-кривых для базовых моделей с лагами от 1 до 8 кварталов

3.4. Анализ нелинейностей по финансовым переменным

Статистическая проверка экономических соображений. Исходя из интуитивных экономических соображений, высказанных в подразд. 3.1, оценим влияние нелинейности по двум переменным, характеризующим отношения соответственно балансовой прибыли br_sa и собственного капитала sk_sa к чистым активам, включив в модель полиномы по данным переменным, чтобы оценить степень нелинейностей, которые потенциально могут встречаться в моделях вероятности дефолта. В дальнейшем возможно рассмотрение нелинейных эффектов и по другим переменным, например, характеризующих размер банка.

В качестве основного инструмента для выбора статистически правильной степени будет использоваться тест максимального правдоподобия (LR-тест)⁶. Нулевая гипотеза данного теста формулируется

⁶ LR-тест применим для сравнения качества двух моделей ($m1$ и $m2$), оцененных методом максимального правдоподобия.

К сравниваемым моделям предъявляется требование включения: « $m1$ nested in $m2$ » – первая ограниченная (restricted) может быть получена из второй расширенной (unrestricted) путем наложения ограничения на вторую, в частности, о равенстве коэффициентов при некоторых переменных во второй модели нулю. При этом проверяется гипотеза: соответствующие коэффициенты равны нулю, т.е. ограниченная модель (restricted model) предпочтительнее.

следующим образом: присутствует квадратичная зависимость по переменным bp_ca и sk_ca , а альтернативная гипотеза состоит в том, что присутствует зависимость более высокой степени n одновременно по обоим переменным, $n \leq 8$. Более высокие степени не рассматривались, так как не вносят значительных изменений в форму моделируемой зависимости.

Данный тест статистически подтвердил, что модели с включением полиномов второй степени по переменным bp_ca и sk_ca вполне достаточны и модели с более высокими степенями не дают положительного эффекта (рис. 12). Также не дает дополнительного эффекта использование перекрестных переменных как произведения переисчисленных выше.

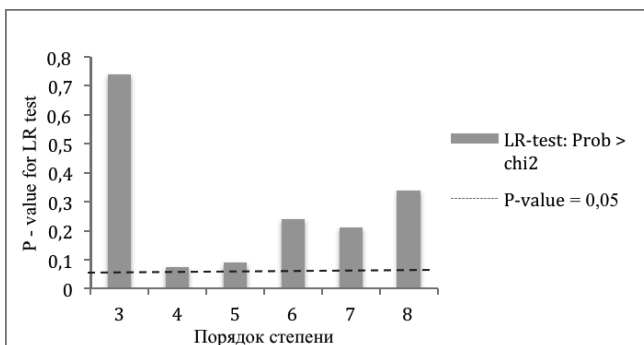


Рис. 12. Результаты LR-теста на включение в модель полиномов более высоких степеней по переменным bp_ca и sk_ca . Сравнение велось с моделью, включающей полиномы второй степени

Модель с финансовыми переменными bp_ca и sk_ca с нелинейностью второго порядка будет называться *моделью с учетом нелинейностей I-go tuna* (basic_exponent1 или b_e1). В результате модель примет вид:

$$P(\text{default} = 1) = \Lambda(sk_ca_{lag2}; (sk_ca_{lag2})^2; \ln_ca_{lag2}; pzs_ke_{lag2}; \ln_oks_ca_{lag2}; bp_ca_{lag2}; (bp_ca_{lag2})^2; ncb_ca_{lag2}).$$

Эконометрический подход к анализу нелинейностей. При этом подходе проводится дополнительный анализ нелинейностей, включаю-

ший использование квадратов и кубов объясняющих переменных в базовой эконометрической модели Basic за исключением переменных размера банка \ln_ca_lag2 и качества управления $\ln_oks_ca_lag2$. Более высокие степени переменных не рассматривались, чтобы исключить влияние высокой мультиколлинеарности на значимость коэффициентов. При таком подходе для проверки целесообразности включения нелинейности будем последовательно отбрасывать незначимые объясняющие финансовые переменные, начиная с наименее значимых, одновременно анализируя статистические характеристики модели.

Данный алгоритм привел к результату, близкому ранее полученному, но по статистическим характеристикам эта модель превзошла первую. Данная модель будет называться *моделью с учетом нелинейностей 2-го тупа* ($basic_exponent2$ или b_e2):

$$P(\text{default} = 1) = \Lambda(sk_ca_{lag2}; (sk_ca_{lag2})^2; \ln_ca_{lag2}; pzs_ke_{lag2}; \ln_oks_ca_{lag2}; bp_ca_{lag2}; (bp_ca_{lag2})^2; ncb_ca_{lag2}; (ncb_ca_{lag2})^2),$$

и именно ее улучшением мы будем заниматься в последующей части работы.

Анализ включения нелинейности по размеру банка. Анализ статистических характеристик полученных ранее моделей показывает, что коэффициент при показателе размера банка, выраженном в виде логарифма активов \ln_ca (\ln_ca_lag2), далеко не всегда значим. В то же время как существующая практика анализа статистических данных по дефолтам российских банков, так и экономические соображения, приведенные в подразд. 3.1, показывают, что такая зависимость существует. Вполне вероятно, что зависимость носит U-образный характер и неудовлетворительно описывается линейной зависимостью. Эту гипотезу мы попытаемся подтвердить для итоговой модели.

3.5. Анализ фактора времени

Одним из недостатков простой логистической модели является неучет фактора времени.

Учет фактора времени по годам. Проверим, значим ли фактор времени, введя дамми-переменные на 1998–2009 гг. Убедившись в значимости группы переменных, отвечающих за учет фактора времени, при помощи LR-теста, мы приступили к последовательному исключению индивидуальных дамми-переменных, начиная с наименее значимых и наблюдая за статистическими характеристиками моделей.

Выполнение этого алгоритма привело к тому, что в модели осталась только одна дамми-переменная, на 2009 г. Вполне возможно, это связано с трансформацией требований к банкам в условиях кризиса 2008 г.

Учет сезонности. Подтвердив наличие проблемы сезонности при помощи LR-теста, мы ввели дамми-переменные на кварталы. Далее мы следовали ранее описанному алгоритму по исключению незначимых переменных. В результате в модель была включена дамми-переменная на 1 квартал. Объяснением может служить тот факт, что многие банки предпочитают более консервативную отчетную политику в первом квартале года, страхуясь от возможных макроэкономических и индивидуальных рисков.

Полученная модель будет называться *моделью с дополнительным учетом фактора времени* (basic_exponent2_time или b_e2_t).

В результате анализа фактора времени модель приняла следующий вид:

$$P(\text{default} = 1) = \Lambda(sk_ca_{lag2}; (sk_ca_{lag2})^2; \ln_ca_{lag2}; pzs_ke_{lag2}; \ln_oks_ca_{lag2}; bp_ca_{lag2}; (bp_ca_{lag2})^2; ncb_ca_{lag2}; (ncb_ca_{lag2})^2; d_{09}; d_{q1}),$$

а ее статистические характеристики приведены далее в табл. 7 одновременно с другими моделями.

3.6. Учет макроэкономических и институциональных факторов

Учет особенностей внешней среды, в которой функционирует банк (макроэкономических и институциональных факторов) позволяет более точно прогнозировать вероятности дефолта банков. Поэтому далее рассматриваются соответствующие группы переменных, выбор наиболее значимых с эконометрической точки зрения и использование их в моделях.

Выбор макроэкономических переменных. Использование макропеременных потенциально позволяет улучшить прогнозное качество модели. При этом во внимание принимаются не столько абсолютные значения макропеременных (например, валютный курс рубля к доллару), сколько относительные характеристики, отражающие их внезапные и резкие изменения, которые оказывают влияние на оперативную деятельность банка. Поэтому ряд используемых далее объясняющих переменных учитывал темпы роста макроэкономических переменных.

При сборе данных были использованы ресурсы Единого архива экономических и социологических данных, Госкомстата России, а также Центрального банка РФ. Изначально были отобраны шесть макроэкономических переменных, их описание приведено в табл. 5. Для этих показателей была рассчитана поквартальная динамика за период с 1998 г. по 2011 г.

При выборе макроэкономических переменных учитывалось, что они могут сильно коррелировать между собой, что может привести в модель проблему мультиколлинеарности с вытекающими из этого последствиями. Поэтому был проведен первоначальный отбор переменных, в результате которого мы остановились на двух из них: ежеквартальные темпы роста ВВП GDP_gr и индекс потребительских цен CPI.

Таблица 5. Макроэкономические переменные, которые рассматривались как кандидаты на включение в модель вероятности дефолта кредитных организаций

Показатель	Наименование и описание
exp_imp	Отношение экспорта к импорту за квартал
gdp_gr	Индекс реального ВВП с поправкой на сезонность. Использовался темп прироста относительно предыдущего квартала
cpi	Индекс потребительских цен. Отражает темп роста цен за квартал
exrate_gr	Средний курс доллара на ММВБ за квартал. Использовался темп прироста (в %) текущего значения переменной относительно предыдущего квартала
unemp	Норма безработицы (в среднем за квартал)
incom_gr	Индекс реальных денежных доходов населения с поправкой на сезонность. Использовался темп прироста относительно предыдущего квартала

Выбор величины лага по макроэкономическим переменным производился так же, как и по финансовым переменным. В результате оптимальной оказалась величина лага в два квартала. Разумность включения макроэкономических переменных была подтверждена LR-тестом.

Данная модель будет называться *моделью с дополнительным учетом фактора времени и макроэкономических переменных* (basic_exponent2_time_macro или b_e2_t_m). В результате модель приняла вид:

$$P(\text{default} = 1) = \Lambda(sk_ca_{lag2}; (sk_ca_{lag2})^2; ln_ca_{lag2}; pzs_ke_{lag2}; ln_oks_ca_{lag2}; bp_ca_{lag2}; (bp_ca_{lag2})^2; ncb_ca_{lag2}; (ncb_ca_{lag2})^2; d_{09}; d_{q1}; gdp_gr_{lag2}; cpi_{lag2}).$$

Учет институциональных переменных. Учет институциональных переменных, проведенный ниже, принимает во внимание не только типовые факторы институционального характера, но и особенности их реализации в России. В модели предлагается рассмотреть возможность использования пяти институциональных переменных, перечисленных в табл. 6.

Таблица 6. Описание институциональных факторов: территориальное положение *region*, участие банка в ССВ *dia*, индекс Лернера

Переменная	Значение	Наименование и описание
region	0	Головной офис зарегистрирован не в Москве
	1	Московский банк. Головной офис зарегистрирован в Москве
dia	0	Банк принимает участие в ССВ
	1	Банк не входит в ССВ
l_index	Индекс Лернера: отражает уровень монопольной власти отдельно взятого банка	

Информация о территориальном расположении головного офиса банков доступна на сайте Банка России. Реестр банков – участников

системы обязательного страхования вкладов был подготовлен экспертами информационно-правового портала «Гарант». Индекс Лернера был рассчитан нами в соответствии с методологией, предложенной в статье (Fungacova, Weill, 2009).

Анализ влияния участия банка в системе страхования вкладов. После включения переменной *dia* в модель обнаружилось, что она незначима, и статистические свойства модели ухудшились. Возможно, причина в том, что в ССВ входят слишком разнообразные банки, в результате чего эффект членства в системе размывается. По сути, требования вхождения в систему уменьшают регуляторную роль рыночной дисциплины.

Адекватность включения в модель оставшихся институциональных переменных *l_index* и *region* была подтверждена LR-тестом.

В результате *модель с дополнительным учетом фактора времени, макроэкономических и институциональных переменных* (*basic_exponent2_time_macro_institutional* или *b_e2_t_m_i*) включила в себя две институциональные переменные и приняла следующий вид:

$$P(\text{default} = 1) = \Lambda(sk_ca_{lag2}; (sk_ca_{lag2})^2; \ln_ca_{lag2}; pzs_ke_{lag2}; \ln_oks_ca_{lag2}; bp_ca_{lag2}; (bp_ca_{lag2})^2; ncb_ca_{lag2}; (ncb_ca_{lag2})^2; d_{09}; d_{q1}; gdp_gr_{lag2}; cpi_{lag2}; l_index; region).$$

Для проверки справедливости сформулированного ранее утверждения о нелинейности зависимости вероятности дефолта относительно размера банка, включим в модель вторую степень по данной переменной. Эта модель будет называться *моделью с дополнительным учетом фактора времени, макроэкономических и институциональных факторов, квадратичностью размера банка* (*basic_exponent2_time_macro_institutional_size* или *b_e2_t_m_i_s*).

$$P(\text{default} = 1) = \Lambda(sk_ca_{lag2}; (sk_ca_{lag2})^2; \ln_ca_{lag2}; (\ln_ca_{lag2})^2; pzs_ke_{lag2}; \ln_oks_ca_{lag2}; bp_ca_{lag2}; (bp_ca_{lag2})^2; ncb_ca_{lag2}; (ncb_ca_{lag2})^2; d_{09}; d_{q1}; gdp_gr_{lag2}; cpi_{lag2}; l_index; region).$$

Данная модель будет считаться *финальной моделью*. В нее входит ряд незначимых переменных, включение которых экономически обоснованно и соответствует проверяемым гипотезам. Дело в том, что при практическом применении моделей данного класса следует различать теоретические и прикладные результаты. Итоговая таблица, позволяющая сравнить предложенные модели приведена ниже (табл. 7), а соотношение соответствующих ROC-кривых – на рис. 13.

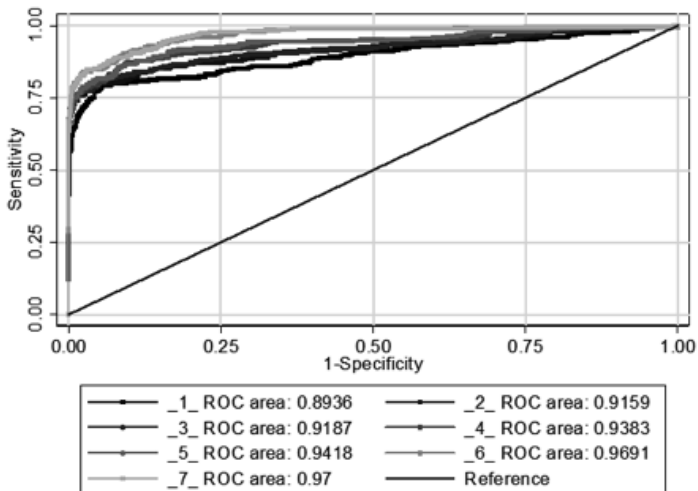


Рис. 13. Сравнение ROC-кривых для ключевых моделей

Таблица 7. Итоговая таблица предложенных моделей. Значимость коэффициентов в построенных моделях. Сравнение статистических характеристик построенных моделей

Название модели	Базовая с учетом нелинейн. 1-го типа	Базовая с учетом нелинейн. 2-го типа	С доп. уч. фактора времени	С доп. уч. фактора времени, макроэ. перемен.	С доп. уч. фактора времени, макроэ. и институт. перемен.	С доп. уч. фактора времени, макроэ. перемен., квадр. размера
Обозначение модели	b_e1_2_	b_e2_3_	b_e2_t_4_	b_e2_t_m_5_	b_e2_t_m_i_6_	b_e2_t_m_i_s_7_
Значимость коэффициентов в построенных моделях						
sk_ca_lag2	-11.05***	-9.75***	-9.98***	-9.10***	-12.43***	-10.91***
(sk_ca_lag2) ²	16.56***	14.56***	15.04***	14.39***	16.81***	14.09***
ln_ca_lag2	-0.01	0.008	-0.13*	-0.10***	-0.11	-1.94**
(ln_ca_lag2) ²						.06**
bp_ca_lag2	-71.90***	-74.52***	-72.42***	67.51***	-61.50***	-60***
(bp_ca_lag2) ²	1014***	976***	1073***	976***	1088***	1039***
nbc_ca_lag2	4.47***	-6.20***	-4.80*	-4.95*	-4.25***	-4.55*
(nbc_ca_lag2) ²		24.05***	22.28***	22.50***	15.54**	16.08**
pzs_ke_lag2	4.72***	4.33***	3.94***	4.23***	5.17	4.96***
ln(oks_ca_lag2)	-1.09***	-1.08***	-1.08***	-1.01***	-1.19	-1.18***
d_09		1.61***		2.15***	2.21	2.26***

Название модели	Базовая	Базовая с учетом нелинейн. 1-го типа	Базовая с учетом нелинейн. 2-го типа	С доп. уч. фактора времени	С доп. уч. фактора времени, макроэ. перемен.	С доп. уч. фактора времени, макроэ. и инстигуц. перемен.	С доп. уч. фактора времени, макроэ. и инстигуц. перемен., квадр. размера
Обозначение модели	basic _1_	b_e1 _2_	b_e2 _3_	b_e2_t _4_	b_e2_t_m _5_	b_e2_t_m_i _6_	b_e2_t_m_i_s _7_
d_q1				-1.07***	-1.28***	-1.41***	-1.47***
gdp_gr_lag2				.17**		0.11	.12
cri_lag2				.10***		.10***	.10***
l_index						-2.51***	-2.38***
region						2.85***	2.91***
Сравнение статистических характеристик построенных моделей							
Критерий сравнения							
Pseudo R ²	0.5219	0.59	0.6046	0.6279	0.6403	0.7058	0.7095
S под ROC	0.8936	0.9159	0.9187	0.9383	0.9418	0.9691	0.97
Sensitivity	72.30%	75.90%	77.34%	78.42%	79.14%	83.81%	84.53%
Specificity	97.20%	97.68%	98.16%	96.64%	96.96%	97.20%	97.04%
Pt (верное предсказ.)	92.67%	93.72%	94.37%	93.32%	93.72%	94.76%	94.76%
Pt(ExWork)	27.7%	24.10%	22.66%	21.58%	20.86%	16.19%	15.47%

Примечания. *** — коэффициент значим на уровне 1%, ** — коэффициент значим на уровне 5%; * — коэффициент значим на уровне 10%.

3.7. Тестирование качества построенной модели

Тестирование на влияние пропусков. В исходной базе данных присутствовало достаточно много пропущенных значений по ключевым финансовым показателям.

В силу этого был реализован алгоритм статистической проверки влияния пропусков в исходной модели на полученные результаты, предложенный в (Магнус и др., 2007). Основная идея данного алгоритма – ввести статистику от дамми-переменных на пропущенные значения по объясняющим переменным, а затем проверить ее значимость. Сам алгоритм состоит из следующих шагов:

Шаг 1. Для каждой объясняющей переменной вводится своя дамми-переменная, которая принимает значение «1», если есть пропуск по выбранному финансовому показателю и «0» – в иных случаях.

Шаг 2. Находится сумма дамми-переменных по всем банкам для выбранного финансового показателя.

Шаг 3. Подстановка результата шага 2 в модель и получение оценки.

Проведение данного теста показало, что в финальной спецификации модели отсутствует статистически значимое влияние пропущенных значений на результаты оценок коэффициентов регрессии, так как все введенные дамми-переменные оказались незначимыми на всех разумных уровнях значимости.

Тестирование на переобучаемость. Возможным недостатком построенной логистической регрессии является тенденция к переобучаемости, т.е. высокий уровень зависимости полученных оценок и статистических свойств модели от исходной выборки. В данной работе получил развитие алгоритм перемешивания, описанный в (Hosmer, Lemeshow, 2010).

Смысл базового алгоритма сводится к следующему:

1. Фиксируется число обанкротившихся банков на максимально возможном уровне (для нашей выборки после формирования лагов – 365 случаев).

2. Случайным образом выбирается 1000 банков из числа успешно функционирующих.

3. Оценивается регрессия с использованием выбранных и дефолтных банков, сохраняются знаки при полученных оценках коэффициентов.

4. Шаги 1–3 повторяются 1000 раз, проверяется стабильность знаков.

Предложено дополнить алгоритм перемешивания тем, что на третьем этапе сохраняли не только знаки, но и уровень значимости коэффициентов.

Результаты тестирования представлены в табл. 8.

Таблица 8. Тестирование на переобучаемость. Показаны результаты оценивания 1000 подвыборок с позиции знаков и значимости коэффициентов

Переменная	Знак «+»	Знак «-»	***	**	*	Незначим
sk_ca_lag2	0	1000	599	396	5	0
(sk_ca_lag2) ²	1000	0	188	785	27	0
lnca_lag2	0	1000	35	676	289	
(lnca_lag2) ²	1000	0	134	527	336	3
bp_ca_lag2	0	1000	1000	0	0	0
(bp_ca_lag2) ²	1000	0	205	623	169	3
ncb_ca_lag2	1000	0	2	83	670	245
(ncb_ca_lag2) ²	1000	0	0	1	625	374
pzs_ke_lag2	1000	0	700	213	86	1
ln_oks_ca_lag2	1000	0	1000	0	0	0
d_09	1000	0	1000	0	0	0
d_q1	0	1000	119	667	187	27
gdp_gr_lag2	1000	0	0	0	11	989
cpi_lag2	1000	0	46	937	17	0
l_index	0	1000	121	634	245	0
region	1000	0	1000	0	0	0

Примечания. *** – коэффициент значим на уровне 1%; ** – коэффициент значим на уровне 5%; * – коэффициент значим на уровне 10%.

Интерпретация результатов достаточно очевидная. С точки зрения сохранения экономического смысла интерес представляют знаки при полученных оценках вне зависимости от исходной выборки. Знаки коэффициентов при всех переменных абсолютно стабильны, а следовательно, модель будет улавливать тренд, а не зависеть от исходной выборки. Учитывая значимость большинства коэффициентов при переменных, можно говорить о том, что модель не подвержена эффекту переобучаемости.

3.8. Интерпретация результатов

Прежде чем приступить к непосредственной интерпретации результатов, необходимо произвести обобщение коэффициентов при переменных для 1000 моделей, построенных по всем подвыборкам⁷. Итоговые коэффициенты приравнивались к арифметическому среднему, рассчитанному по коэффициентам при соответствующих переменных у 1000 моделей.

В результате модель приняла вид

$$\begin{aligned}
 P(\text{default} = 1) = & \\
 & \Lambda(-11.0584 \cdot sk_ca_{lag2} + 14.1532 \cdot (sk_ca_{lag2})^2 - 1.9416 \cdot \\
 & \ln_ca_{lag2} + 0.0674 \cdot (\ln_ca_{lag2})^2 + 4,6421 \cdot pzs_ca_{lag2} - 1,0783 \\
 & \cdot \ln_oks_ca_{lag2} - 61.6552bp_ca_{lag2} + 1053 \cdot (bp_ca_{lag2})^2 \\
 & - 4,7581 \cdot ncb_ca_{lag2} + 17,0004 \cdot (ncb_ca_{lag2})^2 + 2,2343 \cdot d_{09} \\
 & - 1,4817 \cdot d_{q1} + 1,1198 \cdot gdp_ca_{lag2} + 0,1023 \cdot cpi_{lag2} - 2,35 \cdot l_{index} \\
 & + 2,96 \cdot region + 1,67).
 \end{aligned}$$

Следует отметить, что направление влияния изменения фактора x на вероятность выживания банка может быть определено по знаку:

$$\frac{\partial \Lambda(f(x))}{\partial x_i} \cdot \Delta x = \lambda(f(x)) \cdot f'(x) \cdot \Delta x, \text{ где множитель } \lambda(f(x)) \text{ всегда по-}$$

⁷ Как было сказано ранее, получение финальной спецификации на этапе построения модели осуществлялось по одной из 1000 подвыборок.

ложителен. Иными словами, характер влияния роста значения переменной на вероятность дефолта определяется по знаку $f'(x)$ (в случае линейной зависимости – по знаку коэффициента при переменной x) (Магнус и др., 2007).

Анализ влияния финансовых переменных. Зависимость вероятности дефолта банка *от доли собственного капитала банка относительно его чистых активов sk_ca* соответствует нашим ожиданиям. Эта зависимость, как мы и предполагали, нелинейна (рис. 14). В качестве оптимального значения по результатам проведенного исследования можно считать значение отношения, равное 0,4.

При слишком высоком отношении собственного капитала к чистым активам (более 0,4) дальнейшее увеличение рассматриваемого показателя сигнализирует о росте вероятности дефолта кредитной организации.

При умеренных значениях sk_ca (менее 0,4) рост показателя уменьшает вероятность дефолта КО, формируя «подушку безопасности» банка.

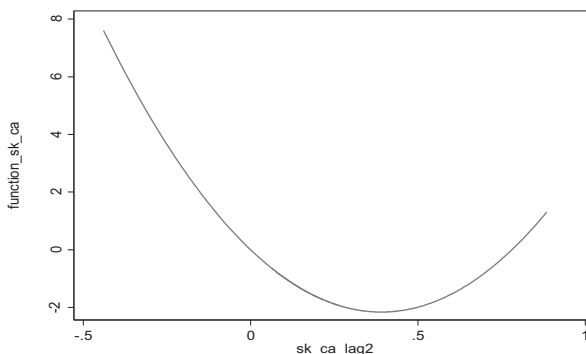


Рис. 14. Влияние отношения sk_ca на вероятность дефолта банка:
 $f(sk_ca) = -11.0584 \cdot sk_ca_lag2 + 14.1532 \cdot (sk_ca_lag2)^2$

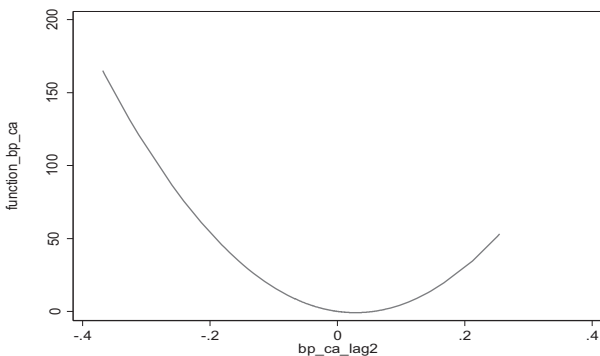


Рис. 15. Влияние отношения bp_ca на вероятность дефолта банка:
 $f(bp_ca) = -60 \cdot bp_ca_lag2 + 1053 \cdot (bp_ca_lag2)^2$

Особенно ярко это проявляется при экстремально низких значениях показателя.

Зависимость вероятности дефолта банка от доли балансовой прибыли банка относительно его чистых активов bp_ca полностью соответствует нашим предсказаниям о влиянии данной переменной на вероятность выживания банка: связь и для данной переменной нелинейна (рис. 15). Оптимальные значения лежат в районе 0,05.

При слишком высоком отношении балансовой прибыли к чистым активам (более 0,05) дальнейшее увеличение рассматриваемого показателя приводит к росту вероятности дефолта кредитной организации. На эффективном рынке получение сверхприбылей сопряжено с огромными рисками, которые, по-видимому, принимает на себя достигающий их банк.

При умеренных значениях рост показателя увеличивает вероятность выживания КО, bp_ca (менее 0,05). Особенно ярко это проявляется при очень низких значениях показателя (менее 0). Ведь прибыль, как уже было сказано, основной источник средств для развития в успешной организации.

Доля негосударственных ценных бумаг в активах банка связана нелинейной связью с вероятностью выживания банка. Для банков, вкладывающих умеренное количество денежных средств в негосударственные ценные бумаги (менее 10%), отсутствует негативное влияние данного показателя на вероятность дефолта (рис. 16). Такие вложения,

совершаются, как правило, из соображений управления ликвидностью, преимущественно в высоконадежные корпоративные бумаги, доходность по которым выше, чем по государственным.

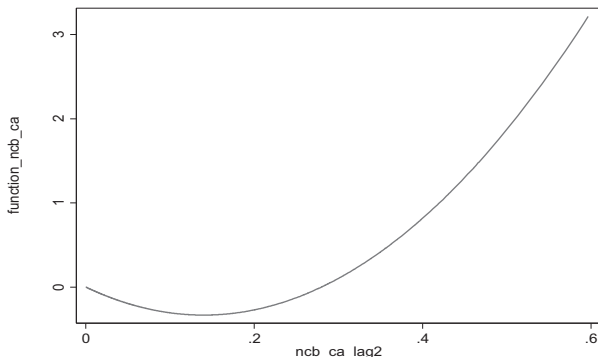


Рис. 17. Влияние отношения \ln_ca на вероятность дефолта банка:

$$f(ncb_ca) = -4,7581 \cdot ncb_ca_{lag2} + 17,0004 \cdot (ncb_ca_{lag2})^2$$

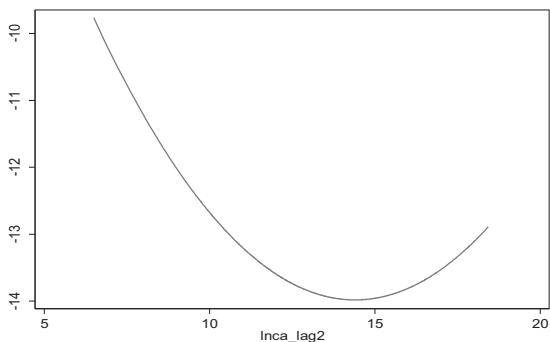


Рис. 16. Влияние отношения ncb_ca на вероятность дефолта банка:

$$f(\ln_ca) = -1.9416 \cdot \ln_ca_{lag2} + 0.0674 \cdot (\ln_ca_{lag2})^2$$

Для банков с крупными вложениями в негосударственные ценные бумаги наблюдается прямая связь данного показателя с ростом вероятности дефолта банка. Некоторые события воспринимаются финансовыми рынками очень болезненно, а это может единовре-

менно принести огромные убытки активно работающему с ценными бумагами (особенно с акциями) банку.

Согласно результатам проведенного анализа, *размер активов банка*, включенный линейно, вопреки нашим ожиданиям не влияет на вероятность его дефолта. В то же время при включении в модель квадратичной объясняющей переменной (модель с учетом квадратичности размера банка) мы наблюдаем опровержение гипотезы о наличии проблемы “Too big to fail” в российской действительности. Конечно, для крупнейших банков на вероятность дефолта и профиль риска банковской деятельности влияет не только размер, но их структура и стратегия развития, которой придерживается банк.

В соответствии с нашими ожиданиями коэффициент при переменной *pzs_ke* принял отрицательное значение. Выдача ссуд ненадежным заемщикам, по причине агрессивной кредитной политики или неадекватной оценки рисков, в результате приводит к повышению вероятности дефолта банка.

Снижение показателя *ln_oks_ca* приводит к повышению вероятности дефолта банка. Как уже было сказано ранее, падающее отношение оборотов по корреспондентским счетам к активам банка за период времени действительно сигнализирует о потенциальных проблемах с проведением платежей или о снижении активности банка.

Анализ влияния фактора времени. Незначимость фактора времени в годовом разрезе говорила бы о том, что модель действительно способна предсказывать кризисы в банковском секторе (через увеличение вероятности дефолта банков), выступать совершенной мерой риска.

Значимой оказалась только переменная *d_09* для 2009 г. Коэффициент при данной переменной положителен, т.е. в 2009 г. вероятность дефолта у банков была выше (произошла недооценка риска). Это естественно: в это время мировую экономику потряс финансовый кризис.

При этом дамми-переменная на 1999 г. (кризис 1998–1999 гг.) и на 2004 г. (банковский кризис 2004 г.) не оказались значимыми. Это говорит о том, что модель предсказывает относительно неглубокие финансовые кризисы, но требуется дополнительная корректировка модели в условиях системного спада.

По-видимому, важен сам механизм распространения кризисных явлений, ведь кризис 1998–1999 гг. никак нельзя назвать неглубоким.

Если предположить, что отчетность банков в 2008–2009 гг. не была существенно искажена, то возможная причина появления дамми-переменной заключается в том, что модель в недостаточной степени учла предкризисную ситуацию до 1998 г., что крайне важно для понимания различий из-за низкого качества соответствующих балансовых отчетных данных. Учет такого канала влияния, как финансирование нерезидентами или иные ненаблюдаемые факторы, интересен для развития исследования, но на сегодняшний день релевантные данные приемлемого качества отсутствуют.

Согласно результатам построенной модели, банки допускают дефолт в 1 квартале каждого года с меньшей вероятностью (знак переменной d_q1). Возможно, это связано с колебаниями деловой активности в банковском секторе, а также с особенностями подготовки финансовой отчетности, отмеченной ранее, и пруденциального надзора, усиливающегося в середине и к концу года.

Анализ влияния макроэкономических переменных. Переменная gdp_gr , отражающая поквартальные темпы прироста ВВП, оказалась незначимой, что достаточно неожиданно. Возможно, что причина такого результата кроется в использовании поквартальных темпов прироста, а не годовых. Потенциально это могло размыть влияние данного фактора на вероятность дефолта банка. Кроме того, положительное влияние высоких темпов экономического роста могло быть учтено неявно, через соответствующее влияние на финансовые переменные. Наличие второй макропеременной также может размывать влияние gdp_gr на вероятность дефолта КО (Карминский, 2005).

В модель включен индекс потребительских цен cpi (прокси для показателя уровня инфляции). Оценка коэффициента при данном показателе положительна. Это означает, что с увеличением уровня инфляции вероятность дефолта банка растет. Одним из объяснений может служить то, что увеличение уровня цен (инфляции) снижает реальную доходность банка по выданным ранее кредитам. При этом вкладчики имеют возможность забрать ранее вложенные в банк средства, а затем повторно разместить их под более высокий процент, что снижает маржу банка в условиях кризисных явлений.

Анализ влияния институциональных переменных. Высоким значениям индекса Лернера l_index , которые соответствуют значительной монопольной власти банка, соответствует пониженная вероятность

дефолта. Это согласуется с выводами статьи (Fungacova, Weill, 2009) и не противоречит здравому смыслу: *ceteris paribus*, в условиях низкой конкуренции вероятность выживания каждого отдельного банка выше. Последний факт связан с тем, что увеличение конкуренции вынуждает банк брать на баланс больше рисков, тем самым повышая вероятность дефолта.

Согласно полученным результатам, коэффициент при дамми-переменной *region* отрицателен. Это означает, что с позиции надежности для московских банков издержки агрессивной конкурентной среды превышают выгоды от развитой инфраструктуры и качественных человеческих ресурсов: в столице вероятность выживания банков ниже. Возможно, это связано с более консервативным отношением Банка России к отзыву лицензий у региональных банков, чтобы не ослаблять и без того не очень высокую конкуренцию на многих местных рынках.

3.9. Тестирование предсказательной силы модели. Прогноз вне выборки

Прогноз вне выборки проводился по данным за 2010–2011 гг., которые не использовались при построении моделей.

Для определения качества прогноза были использованы следующие показатели:

- количество предсказанных дефолтов в 2010–2011 гг.;
- среднеквартальная величина группы риска в 2010–2011 гг.

Таблица 8. Тестирование предсказательной силы модели. Количество верно предсказанных дефолтов в зависимости от определения группы риска

Критерий причисления банка к группе риска: вероятность дефолта банка $\geq X$	Величина группы риска, среднеквартальная	Количество (доля) верно предсказанных дефолтов, количество, (%)
$X = 10\%$	54	16 (84%)
$X = 20\%$	34	12 (63%)
$X = 30\%$	30	12 (63%)
$X = 40\%$	28	10 (52%)

Всего за 2010–2011 гг. было зарегистрировано 19 случаев дефолта. По нашему мнению, к группе риска следует причислять те банки, вероятность дефолта которых превышает 30%. Это позволяет одновременно сократить размер группы риска и предсказать значительное число дефолтов – 63% (табл. 8).

По нашему мнению, это весьма убедительный результат. При этом на практике определение параметра X зависит от того, насколько жестко подходит пользователь модели (например, ЦБ РФ) к дистанционному мониторингу деятельности банков и наличию ресурсов для проведения детального (например, выездного) мониторинга попавших в группу риска банков.

На практике для повышения точности оценивания и выявления банков, требующих повышенного внимания, можно использовать одновременно альтернативные модели, например, на основе рейтингов, а также рассмотренные в следующем разделе.

4. Сравнение альтернативных моделей вероятности дефолта

В данном разделе мы сравним полученную модель с альтернативными. В качестве альтернатив рассматривается использование панельных данных, Z -индекса, нейронной сети, а также некоторые другие работы по прогнозированию вероятности дефолта.

4.1. Логистическая модель бинарного выбора с использованием панельных данных

Мы оценили модель финальной спецификации с использованием панельных данных, а затем планировали сравнить ее предсказательную силу с той, которую продемонстрировала простая логистическая модель.

При построении логистической модели с панельными данными всегда возникает вопрос, случайный или фиксированный эффект предпочтительнее использовать.

Фиксированный эффект используется в том случае, когда данные, использованные для построения модели, нельзя рассматривать как

случайную выборку из некоторой более широкой совокупности. Такая интерпретация больше подходит для случаев, когда субъектами исследования являются крупные компании, а это и есть наш случай (Носко, 2005).

После этого был проведен тест Хаусмана на различие в полученных оценках коэффициентов между простой и панельной логистическими моделями. На основании его результатов гипотеза о несистематическом различии в коэффициентах не была отвергнута на всех разумных уровнях значимости. Примечательно, что значимость коэффициентов при переменных, статистические характеристики модели и ее предсказательная сила также сохранились.

Таким образом, использование панельных данных не привело к ощутимому улучшению построенной модели вероятности дефолта банка.

4.2. Нейронная сеть для прогнозирования вероятности дефолта банка

Как и в предыдущих случаях, для тестирования предсказательной силы модели была сформирована выборка (out-of-sample) банков, которые обанкротились в 2010–2011 гг. (всего 19 банков). Далее модель оценивалась на выборке с 1998 г. по 2009 г.

Выбор нейронной сети как основного бенчмарка продиктован тем соображением, что данная модель является обобщением логистической регрессии, и в эмпирической литературе данный класс моделей находит достаточно широкое применение.

На вход подавались финансовые показатели из финальной спецификации. Топология нейронной сети – однослойная нейронная сеть, каждый нейрон которой состоит из логистической активационной функции с неизвестными параметрами. Выход – бинарная переменная, принимающая значение 1 или 0 (1 – банк обанкротился, 0 – банк продолжает операционную деятельность). Процедура обучения⁸ нейронной сети – метод наименьших квадратов. Также на данном этапе были добавлены макроэкономические и институциональные переменные.

⁸ Процедура обучения нейронной сети – процедура нахождения неизвестных параметров данной математической модели.

По результатам тестирования вне выборки нейронной сети удалось предсказать восемь дефолтов (42%) за 2010–2011 гг. Это говорит о том, что нейронная сеть проигрывает построенной ранее модели вероятности дефолта.

4.3. Модель с использованием в качестве объясняющей переменной Z-индекса

Согласно мнению исследователей, увеличение Z-индекса приводит к улучшению финансовой стабильности банка и, как следствие, уменьшению вероятности дефолта.

В нашем же случае включение данной переменной в модель не привело к улучшению ее качества, более того, переменная оказалась незначимой на всех разумных уровнях значимости.

Причины этого следует искать в порядке расчета индекса:

$$Z = \frac{E[ROA] + \frac{SK}{CA}}{\sigma(ROA)}.$$

Данный показатель зависит от трех величин: отношения собственного капитала к чистым активам банка, рентабельности его активов и ее волатильности.

Фактически первые два уже учтены в модели (через переменные `sk_ca` и `bp_ca` соответственно), а третий компонент, по-видимому, не играет доминирующей роли.

Заключение

В данной работе построена адекватная модель вероятности дефолта банка, которая продемонстрировала убедительную предсказательную силу при тестировании вне выборки: при умеренном количестве банков в группе риска было верно предсказано более 60% произошедших в 2010–2011 гг. дефолтов. Это, в свою очередь, подтверждает применимость автоматической кластеризации в совокупности с подходом CAMELS при отборе наилучших объясняющих переменных для построения моделей вероятности дефолта.

Согласно результатам нашей работы, была обнаружена квадратическая зависимость вероятности дефолта банка от ряда относительных финансовых переменных: по капитализации, по прибыльности, по доле негосударственных ценных бумаг в активах банка, а также по размеру банковских активов. Особенно интересен последний результат: без учета нелинейности относительно размеров банка этот фактор оказывается незначимым. Данный вывод может оказаться полезным исследователям, учитывающим влияние размеров банка на различные характеристики его операционной деятельности.

При формировании базы данных для исследования мы столкнулись с проблемами «замусоренности» финансовой статистики, несбалансированности данных, однако нашли пути решения этих проблем. Кроме того, мы улучшили модель при помощи макроэкономических и институциональных характеристик операционной среды банка. Перспективным направлением развития модели видится использование банковских рейтингов для получения более точной оценки вероятности дефолта банка.

Литература

Инструкция ЦБ РФ от 16 января 2004 г. N 110-И «Об обязательных нормативах банков» (с изменениями и дополнениями) (<http://base.garant.ru/584347/>).

Карминский А.М., Пересецкий А.А., Петров А.Е. (2005) Рейтинги в экономике: методология и практика / под ред. А.М. Карминского. М.: Финансы и статистика, 2005.

Магнус Я.Р., Катышев П.К., Пересецкий А.А. (2007) Эконометрика: начальный курс. М.: Дело, 2007.

Пересецкий А.А. (2010) Модели причин отзыва лицензий российских банков. Препринт WP/2010/085. М.: Российская экономическая школа, 2010.

Носко В.П. (2005) Эконометрика для начинающих (дополнительные главы). М.: ИЭПП, 2005.

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

Федеральный закон «О банках и банковской деятельности» от 02.12.1990 N 395–1, ст. 20 (<http://www.consultant.ru/popular/bank/>).

Anzoategui D., Pería M., Melecky M. (2012) Bank competition in Russia: An examination at different levels of aggregation // *Emerging Markets Review*. Vol. 13. No. 1. P. 52–53.

Basel II. Basel Committee on Banking Supervision (2004). *International Convergence of Capital Measurement and Capital Standards. A Revised Framework*. Bank for International Settlements, 2004.

Bluhm C., Overbeck L., Wagner C. (2010) *Introduction to Credit Risk Modeling*. Chapman and Hall/CRC, 2010.

Bock R., Demyanets A. (2012) Bank Asset Quality in Emerging Markets: Determinates and Spillovers. IMF Working Paper WP/12/71. International Monetary Fund, 2012.

Chernykh L., Theodossiou A. (2011) Determinants of Bank Long-Term Lending Behavior: Evidence From Russia // *Multinational Finance Journal*. 2011. No. 15. P. 193–216.

Claeys S., Schoors K. (2007) Bank supervision Russian style: Evidence of conflicts between micro- and macro-prudential concerns // *Journal of Comparative Economics*. 2007. Vol. 35. No. 3. P. 63–657.

Clarke G., Cull R., Shirley M. (2005) Bank privatization in developing countries: A summary of lessons and findings // *Journal of Banking & Finance*. 2005. Vol. 29. No. 8–9. P. 1905–1930.

Demirguc-Kunt A., Huizinga H. (2004). Market Discipline and Deposit Insurance // *Journal of Monetary Economics*. 2004. No. 51. P. 375–399.

Fungacova Z., Solanko L. (2009) Risk-taking by Russian banks: Do location, ownership and size matter? BOFIT Discussion Papers 21/2008. Bank of Finland. Institute for Economies in Transition.

Fungacova Z., Weill L. (2009) How market power influences bank failures: Evidence from Russia. BOFIT Discussion Papers 12/2009. Bank of Finland. Institute for Economies in Transition.

He H., Member, IEEE, Edwardo A. (2009) Learning from Imbalanced Data, *IEEE Transactions on Knowledge and Data Engineering*. Vol. 21. No. 9. P. 1263–1284.

Hosmer D., Lemeshow S. (2000) *Applied logistic regression*. N. Y. (NY): John Wiley and Sons, 2000.

Lanine G., Vennet R. (2006) Failure prediction in the Russian bank sector with logit and trait recognition models // *Expert Systems with Applications*. 2006. Vol. 30. No. 3. P. 463–478.

Mannasoo K., Mayes D. (2009) Explaining bank distress in Eastern European transition economies // *Journal of Banking and Finance*. 2009. Vol. 33 No. 2. P. 244–253.

Micco A., Panizza U., Yanez M. (2007) Bank ownership and performance. Does politics matter? // *Journal of Banking and Finance*. 2007. Vol. 31. No. 1. P. 219–241.

Peresetsky A.A., Karminsky A.M., Golovan S.V. (2011) Probability of default models of Russian banks // *Economic Change and Restructuring*. Vol. 44. No. 4.

Tabak B., Craveiro G., Cajueiro D. (2011) Bank efficiency and Default in Brazil: Causality Tests. Working paper series 253. The Central Bank of Brazil, 2011.

Vernikov A. Government banking in Russia: Magnitude and new features (2011). IWH Discussion Papers. Halle Institute for Economic Research.

Приложения

Приложение 1 Обозначения используемых переменных и их описательные статистики

Название переменной	Количество переменных	Среднее значение	Станд. откл.	Мин. значение	Макс. значение	Пояснение
BP	59270	139057.6	2749515	-7.16e+07	3.18e+08	Балансовая прибыль
CA	59391	1.39e+07	1.90e+08	-6.53e+07	1.02e+10	Чистые активы
CP	58461	102745.2	2242601	-7.16e+07	2.54e+08	Чистая прибыль
DFL	37579	4536610	8.64e+07	0	5.06e+09	Сумма депозитов физических лиц
GDO	32556	1000707	1.66e+07	0	9.12e+08	Долговые обязательства государства и органов местного самоуправления
KE	57608	6798499	1.02e+08	-2082	6.83e+09	Кредиты экономике
KE_F	39454	2093819	3.02e+07	0	1.66e+09	Кредиты физическим лицам и индивидуальным предпринимателям
KE_F_12	36060	1964401	2.99e+07	0	1.58e+09	Кредиты физическим лицам на срок свыше 1 года
KE_Prom	29410	9631169	1.10e+08	0	5.17e+09	Кредиты промышленности
LA	59276	2814943	3.15e+07	0	2.29e+09	Ликвидные активы
NCB	48854	1104642	1.11e+07	0	5.14e+08	Нетоговарственные ценные бумаги
NMO	29726	-799589.2	2.22e+07	-1.37e+09	3.36e+08	Нетто межбанковских операций
OKS	55853	3.53e+07	3.15e+08	0	1.54e+10	Обороты по корреспондентским счетам в ЦБ РФ и других банках
ORCB	57845	146057.9	1839618	0	1.00e+08	Обязательные резервы в ЦБ РФ
OV	59365	4172650	6.65e+07	-8719	4.48e+09	Обязательства до востребования

PZS	47611	369878.6	5511744	0	3.26e+08	Просроченная задолженность по ссудам
PZS_F	22912	194123.4	1609185	0	5.70e+07	Просроченная задолженность по кредитам физическим лицам и индивидуальным предпринимателям
RA	58448	9228285	1.27e+08	0	8.36e+09	Работающие активы
RES	58342	597177.8	9699533	0	7.39e+08	Резервы на возможные потери
SDB	15076	348963.5	2569998	0	1.04e+08	Средства других банков – корреспондентские счета
SK	59629	1596060	2.29e+07	-5.72e+08	1.46e+09	Собственные средства, капитал
SO_LONG	50757	5458812	8.32e+07	0	4.65e+09	Обязательства на срок свыше 1 года
SRTS	41798	2813746	3.14e+07	0	1.84e+09	Средства на расчетных счетах
VBCB	40401	763762.7	5467567	0	2.02e+08	Выпущенные ценные бумаги
VDFL	51810	2813877	6.47e+07	0	4.31e+09	Депозиты физических лиц свыше 30 дней
VDFL_30	54222	551249.4	9284596	0	7.60e+08	Депозиты физических лиц на срок до 30 дней
VDUL	45212	2392271	2.64e+07	0	1.33e+09	Депозиты юридических лиц
NORM_H3	42792	102571.1	3170775	-1000	1.00e+08	Норматив текущей ликвидности
ODB_q	40005	1092664	1.88e+07	-1.34e+08	1.95e+09	Операционные доходы банков
ORB_q	39753	1121457	2.08e+07	-1.34e+08	2.22e+09	Операционные расходы банков
PDFL_q	39671	42527.78	610626.1	-3725184	4.63e+07	Процентные доходы по кредитам физическим лицам
PDK_q	40344	97565.66	1424873	-2.56e+07	1.31e+08	Процентные доходы по кредитам организациям
PDMBK_q	30771	8649.174	116576.5	-976662	9345929	Доходы по кредитам кредитных организаций (КО)
RSA_q	40004	42522.98	552660.7	-471395	4.86e+07	Расходы на содержание аппарата
RUB2_q	36357	832045.2	1.69e+07	-1.39e+08	1.89e+09	Расходы от переоценки счетов в иностранной валюте

Karminsky, A. M., Kostrov, A. V., Murzenkov, T. N. Approaches to Evaluating the Default Probabilities of Russian Banks with Econometric Methods : Working paper WP7/2012/04 [Text] / A. M. Karminsky, A. V. Kostrov, T. N. Murzenkov ; National Research University "Higher School of Economics". – Moscow : Publishing House of the Higher School of Economics, 2012. – 64 p. (in Russian)

Under new Basel accord, improving the probability of default models is one of the key risk-management priorities. The main features of this research are:

- Employing Russian banking statistics for the period from 1998 to 2011.
- Investigating the impact of macroeconomic and institutional factors on default probabilities.
- Testing the accuracy of the developed models.

Also, we have taken into account an influence of the banking market structure and nonlinearities on the estimated probabilities. In addition, we consider the reliability of the developed models.

We have chosen the binary logit – regression with quasi panel data structure and compared it with some alternatives. Our key findings are:

- There is a quadratic relationship between bank's capital adequacy ratio and its probability of default.
- There is a negative relationship between Lerner index and bank's PD.
- Macroeconomic, institutional and time factors significantly improve the model quality.

We believe that these results will be useful for national financial regulatory authorities as well as for commercial banks in risk-management.

Key words: probability of default, PD, banks, Russia, risk-management, IRB approach, Basel II.

JEL: G21, G24, G32.

Препринт WP7/2012/04

Серия WP7

Математические методы анализа решений
в экономике, бизнесе и политике

Карминский А.М., Костров А.В., Мурзенков Т.Н.

**Моделирование вероятности дефолта российских банков
с использованием эконометрических методов**

Зав. редакцией оперативного выпуска *А.В. Заиченко*
Технический редактор *Ю.Н. Петрина*

Отпечатано в типографии
Национального исследовательского университета
«Высшая школа экономики» с представленного оригинал-макета
Формат 60×84 ¹/₁₆. Тираж 60 экз. Уч.-изд. л. 3,95
Усл. печ. л. 3,7. Заказ № . Изд. № 1519
Национальный исследовательский университет
«Высшая школа экономики»
125319, Москва, Кочновский проезд, 3
Типография Национального исследовательского университета
«Высшая школа экономики»