
**МАТЕМАТИЧЕСКИЙ АНАЛИЗ
ЭКОНОМИЧЕСКИХ МОДЕЛЕЙ**

МЕТОДЫ ОЦЕНКИ ВЕРОЯТНОСТИ ДЕФОЛТА БАНКОВ

© 2007 г. **А. А. Пересецкий**
(Москва)

Представлены результаты эконометрического анализа дефолтов российских банков в 1997–2003 гг. Основная цель исследования – выяснить, насколько публично доступная информация балансовых отчетов банков может быть использована для прогнозирования дефолтов банков. Показано, что предварительная экспертная кластеризация банков, а также учет макроокружения повышают качество моделей дефолта. Предложены эвристические критерии оценки качества прогнозной силы моделей. С помощью скользящей регрессии анализируются тенденции развития российской банковской системы после кризиса 1998 г.

1. ВВЕДЕНИЕ И КРАТКИЙ ОБЗОР ЛИТЕРАТУРЫ

В данной работе исследуется возможность применения эконометрических моделей дефолта, использующих публично доступную информацию о балансовых отчетах банков, для прогноза кредитоспособности/устойчивости российских банков.

Многие эксперты предполагают, что такие модели не применимы для российских банков, поскольку кроме финансовых показателей банков и макроэкономического окружения для прогноза устойчивости банка важны многие другие критерии. Например, такие мало формализуемые факторы, как политика, связь банков с промышленными или финансовыми группами, профиль деятельности банка, качество менеджмента.

Кроме того, данные балансовых отчетов банков все еще основаны на российской системе финансовой отчетности, которая значительно менее прозрачна, чем международная система финансовой отчетности. Качество данных также зачастую оставляет желать лучшего. Однако это не является особенностью России. Важность качества балансовых данных в статистическом моделировании банковских рисков была показана на примере банков США в работе (Gunther, Moore, 2003). В работе (Marchesini et al., 2004), где исследованы ошибки в прогнозе высокодоходных облигаций, отмечается, что “балансовые отчеты ... могут быть и бывают серьезно подтасованы”.

Российская банковская система быстро развивалась после начала перехода страны к рыночной экономике. В России много коммерческих банков, хотя их число и сократилось с примерно 2500 в 1995–1996 гг. до 1300 в 2005 г. Это обстоятельство дает богатый материал для построения эконометрических моделей дефолта. Для сравнения заметим, что в (Kolari et al., 2002) для построения моделей используются данные по 1030 крупнейшим банкам США в 1989 г., включая лишь 18 случаев дефолта. В работе (Martin, 1977) рассматривается выборка из 5598 наблюдений по банкам в США за 1970–1976 гг., в которой только 23 дефолта.

Отличие от этих работ рассматриваемая в данной статье выборка существенно более сбалансированная. Известно, что несбалансированные выборки (существенно разное количество нулей и единиц) в случае логистических моделей приводят к смещенным оценкам и потенциально неверным выводам (см., например (Aldrich, Nelson, 1985; Stone, Rasp, 1991)).

Новое Базельское соглашение (Basel-II, 2004) предполагает, что банки могут использовать собственные системы оценки риска (Internal Ratings Based Approach, IRB) для оценки потенциальных партнеров, а органы банковского надзора могут использовать системы раннего предупреждения (Early Warning Systems, EWS) для мониторинга банковской системы. Статистические модели, подобные рассматриваемым в данной работе, могут служить составной частью IRB и EWS. Органы банковского надзора могут оптимизировать усилия, осуществляя дистанционный мониторинг банковской системы с помощью подобных моделей для идентификации потенциально проблемных банков, и затем концентрировать усилия в первую очередь на этих банках для тщательного анализа их состояния.

Такие международные рейтинговые агентства, как Standard and Poor's, Moody's и Fitch публикуют рейтинги надежности многих банков. Эти рейтинги, конечно, могут использоваться для IRB или EWS. Однако их применение в России весьма ограничено. Во-первых, эти наиболее из-

вестные рейтинговые агентства составляют рейтинги лишь для нескольких десятков российских банков. Кроме того, эти рейтинги довольно низкие, и их разброс невелик. Например, в сентябре 2004 г. S&P присваивало рейтинги 22 российским банкам и только трех категорий: CCC, B и BB, причем рейтинг BB был присвоен только Внешторгбанку¹. Во-вторых, рейтинги, присваиваемые рейтинговыми агентствами, довольно консервативны и редко изменяются. В работах (Löffler, 2004) и (Altman, Rijken, 2004) рассматриваются причины стабильности рейтингов и связанной с ней потерей информации. В работе (Soest et al., 2003) показано, что российские рейтинговые агентства в глазах экспертов менее надежны, чем международные рейтинговые агентства².

Существует несколько подходов к статистическому моделированию надежности банков на основе публично доступной информации. Во-первых, для построения эконометрической модели можно использовать рейтинги, присвоенные рейтинговым агентством. Такая модель будет отражать часть информации, содержащейся в рейтинге, которая может быть получена из публично доступной информации. Представляется естественным применение моделей множественного выбора (ordered logit/probit). Когда такая модель построена, ее можно использовать для расчета прогнозных рейтингов любых банков. Прогнозные рейтинги будут отражать мнение экспертов рейтингового агентства. Такой подход был реализован для российских банков в работе (Soest et al., 2003) и для нефинансовых фирм в США – в (Altman, Rijken, 2004).

Второй подход основан на опросе экспертов. Экспертам предлагается присвоить рейтинги некоторому набору реальных или виртуальных банков по выборке финансовых показателей балансовых отчетов. Далее строится эконометрическая модель, отражающая мнения экспертов. Возможное преимущество этого метода состоит в том, что модель строится на основе мнений экспертов, представляющих различные независимые финансовые структуры. Поскольку банки платят агентству за рейтинг, это потенциально может приводить к тому, что рейтинговое агентство не склонно понижать рейтинг. Однако эксперты агентства при составлении рейтинга имеют доступ к исчерпывающей информации о банке. Анализ такого подхода проведен в работе (Soest et al., 2003).

Третий подход, подробный анализ которого и является темой данной работы, состоит в составлении эконометрической модели надежности банка на основе исторических данных о банковских дефолтах. Естественным инструментом здесь являются модели бинарного выбора (logit/probit-модель). Этот подход был впервые применен к российским банкам в работах (Головань и др., 2003, 2004; Peresetsky et al., 2004).

Первым исследованием построения моделей дефолта, насколько нам известно, была работа (Altman, 1968), в которой применялась статистическая модель дискриминантного анализа (DA) для прогноза банкротства нефинансовых фирм. В качестве исходных данных в модели были использованы 5 финансовых индикаторов за один или два года до банкротства (или выживания) фирмы. В (Martin, 1977) впервые модель бинарного выбора была применена к прогнозу дефолтов банков. В этой работе также использовались одно- и двухгодичные горизонты прогноза.

В дальнейшем во многих работах использовались модели бинарного выбора для моделирования вероятности дефолта. В (Wiginton, 1980) показано, что logit-модель дает результаты лучшие, чем дискриминантный анализ, для скоринга потребительских кредитов. В работе (Ohlson, 1980) logit-модель применялась к данным по фирмам за 1970–1976 гг. для выяснения того, какие факторы статистически значимы для прогноза дефолта фирмы через год. В (Westgaard, Wijst, 2001) logit-модель применялась к данным по норвежским фирмам с ограниченной ответственностью за 1995–1999 гг. Показано, что горизонт в два года для прогноза дефолта по данным балансового отчета фирмы является оптимальным. В этой работе в качестве показателя размера фирмы использовался логарифм ее активов. Также было показано, что удаление наблюдений с экстремальными значениями параметров из набора данных улучшает статистическое качество модели. Колари (Kolari et al., 2002) использовал logit-модели для моделирования вероятности дефолта банков США в 1989–1990 гг. Ленокс (Lenox, 1999) на выборке из 949 фирм Великобритании (6416 наблюдений) изучал результаты, полученные по моделям logit, probit и дискриминантного анализа для прогноза дефолта фирм. Он пришел к выводу, что модели logit/probit со спецификацией гетероскедастичности превосходят модели logit/probit без учета гетероскедастичности и модели дискриминантного анализа. Из работы, однако, неясно, связано это улучшение именно с учетом возможной гетероскедастичности или просто с увеличением числа параметров в модели.

¹ На октябрь 2006 г. S&P присвоило рейтинги 25 российским банкам с диапазоном рейтингов CCC, B, BB, BBB; агентство Moody's-Interfax присвоило рейтинги 52 банкам.

² Ситуация может измениться с созданием совместных проектов типа Moody's-Interfax.

Некоторые авторы используют нестатистические методы для моделирования дефолта. В работе (Kolari et al., 2002) применяется вариант алгоритма распознавания образов (TRA, trait recognition algorithm), а в (Espahbodi, Espahbodi, 2003) – рекурсивное разбиение. Другие нестатистические методы включают нейросети, марковские модели, CAMELS и финансовые индикаторы. Однако нет доказательств того, что эти методы дают более надежные результаты, чем статистический подход. Напротив, Альтман (Altman et al., 1994) приходит к заключению, что logit-модели и дискриминантный анализ точнее предсказывают банкротства фирм, чем нейросети. А в работе (Jagtiani, Kolari et al., 2003) авторы приходят к заключению, что “простая линейная logit-модель работает лучше, чем более сложные EWS-модели, например TRA”.

Новизна нашего исследования состоит в следующем. Во-первых, построены модели дефолта российских банков. Во-вторых, рассмотрена необходимость предварительного разбиения банков на кластеры и построения отдельной logit-модели для каждого кластера. Конечно, в этом случае оптимально было бы использовать процедуру кластеризации, ориентированную на наилучшую подгонку logit-модели в каждом кластере. Для этого построена модель, которая совмещает процедуру кластеризации с подгонкой logit-моделей. В-третьих, рассмотрен вопрос о том, до какой степени введение в модель макропеременных может улучшить качество прогноза модели. В-четвертых, предложен новый подход к сравнению моделей. Дело в том, что сравнение моделей только по их статистическим характеристикам или по доле верных прогнозов, – не совсем то, что требуется инвестору. Поэтому предложенный метод основан на эвристических критериях, отражающих дополнительную прибыль инвестора, использующего модель.

2. ПОМОГАЕТ ЛИ КЛАСТЕРИЗАЦИЯ?

Существует лишь небольшое число работ, в которых probit/logit-модели вероятности дефолта строятся на основе исторических данных о дефолтах банков и их балансовых показателях. Данные по банкам США используются в работах (Martin, 1977; Bovenzi et al., 1983; Cole, Gunther, 1995, 1998; Estrella et al., 2000; Kolari et al., 2002); данные по российским банкам – в (Головань и др., 2003, 2004; Peresetsky et al., 2004). В работе (Godlewski, 2004) используются данные по странам с переходной экономикой (исключая Россию). Насколько нам известно, Федеральная резервная система США является единственным органом банковского надзора, использующим в своей работе подобную модель (SEER) как составную часть EWS (Sahajwala, Bergh, 2000).

В данном разделе строятся модели прогноза выживания банка во время финансового кризиса августа 1998 г. Используя эти модели, можно исследовать, какие факторы были наиболее важны для прогноза выживания (или дефолта) банка. До 1998 г. многие банки вместо финансирования реального сектора экономики были вовлечены в спекуляции на финансовых рынках. Это обстоятельство было указано в качестве одной из причин кризиса как правительством, так и ЦБ РФ. Поэтому особое внимание уделяется такому индикатору, как доля кредитов экономике в активах банка.

Российские банки значительно отличаются друг от друга по размеру, направлению деятельности, степени участия на рынке государственных ценных бумаг, доле кредитов экономике в активах, объему депозитов физических лиц и т.п. Многие мелкие банки и более половины из 12 крупнейших банков не выжили во время кризиса 1998 г. Поэтому естественно предположить, что для моделирования вероятности дефолта может быть недостаточно одной logit-модели, и для такого неоднородного набора банков требуется разбить банки на кластеры и строить модели вероятности дефолта отдельно в каждом кластере.

Насколько нам известно, только в работе (Korobow, Stuhr, 1983) была предпринята попытка использовать кластеризацию банков для EWS. В этой работе авторы использовали разбиение на кластеры (peer groups) по размеру банка или по наличию хотя бы одного зарубежного отделения банка.

Насколько улучшает качество прогноза разбиение банков на кластеры? Для ответа на этот вопрос мы ниже используем два метода. Первый – предварительное (экспертное) разбиение банков на кластеры по одному из финансовых показателей. Второй – построение автоматической процедуры кластеризации, ориентированной на построение моделей дефолта в каждом кластере.

Де-факто ЦБ РФ использует кластеризацию в банковском надзоре. Например, для малых и крупных банков устанавливаются разные требования к достаточности капитала (см. табл. 1).

2.1. Финансовые показатели банков. В этом разделе рассматриваются 1595 российских банков и данные их балансов на 1 апреля 1998 г. На основе этой информации строится прогноз того, какие из них потерпели дефолт до 1 апреля 2000 г. Двухлетний период между данными и (воз-

Таблица 1. Требование ЦБ РФ к достаточности капитала банка, %

Период	Собственный капитал более 5 млн. евро	Собственный капитал менее 5 млн. евро
01.02.1999–01.01.2000	8	9
01.01.2000–по настоящее время	10	11

Таблица 2. Финансовые показатели банков

Переменная	Описание переменной
VB	Суммарные активы, валюта баланса
RES	Резервы под возможные потери
KE	Кредиты нефинансовым организациям
GDO	Государственные ценные бумаги
SK	Собственный капитал
LA	Ликвидные активы*
VDFL	Депозиты и счета физических лиц
OS	Основные средства и другие неработающие активы
NDO	Негосударственные ценные бумаги
CA	Чистые активы
BP	Доход до вычета налогов
CFB	Кредиты, полученные от других банков
PNA	Неработающие активы
PZS	Просроченные кредиты (более 5 дней)

* Рассчитано по методологии журнала “Банки и финансы”.

можным) дефолтом выбран по двум причинам: он покрывает среднее время между отзывом лицензии и ликвидацией банка. Кроме того, оказалось, что модели, построенные на двухлетнем периоде, обладают наибольшей прогнозной силой.

Банк отмечается как обанкротившийся, и переменной LIVE присваивается значение 0, если выполняется одно из трех условий:

- лицензия была отозвана до 1 апреля 2000 г.,
- банк на 1 апреля 2000 г. находился под управлением АРКО (Агентство по реструктуризации кредитных организаций),
- банк ликвидирован в результате слияния с другим банком, но перед слиянием его финансовое положение было близко к дефолту (каждый такой случай рассматривался отдельно).

Для всех других банков значение переменной LIVE равно 1. Выборка содержит 263 обанкротившихся и 1306 работающих банков. В нее включены все банки по состоянию на 1 апреля 1998 г. за исключением трех государственных банков – Сбербанк, Внешторгбанк, Внешэкономбанк; также из выборки удалены несколько банков с неполной или ошибочной информацией из балансовых отчетов.

Около 30 финансовых показателей банков рассматривались на этапе выбора моделей. В табл. 2 приведены описания тех из них, которые были значимыми в одной из моделей, описанных в разд. 2, 3 и 4.

Модели, приведенные ниже, включают не сами финансовые показатели банков, а их долю в суммарных активах VB (валюта баланса), т.е. RES/VB, KE/VB, GDO/VB, SK/VB, LA/VB, VD-FL/VB, OS/VB, NDO/VB. Эти индикаторы характеризуют степень участия банка в том или другом виде деятельности. В качестве размера банка лучше всего подходит логарифм валюты баланса. Эти же индикаторы использовались в работах (Головань и др. 2003, 2004; Peresetsky et al., 2004). Аналогичные показатели используются в работах (Martin, 1977; Kolari, 2002; Estrella et al., 2000). В нескольких работах (например (Altman, 2004)) в качестве размера банка также используется логарифм суммарных активов.

Описательная статистика параметров набора банков на апрель 1998 г. приведена в табл. 3.

Таблица 3. Статистика финансовых индикаторов банков

Значения	LNVB	SK/VB	KE/VB	GDO/VB	LA/VB	NDO/VB	OS/VB	VDFL/VB	RES/VB
Среднее	10.72	0.28	0.29	0.07	0.14	0.12	0.20	0.06	0.03
Максимум	17.88	0.99	0.97	0.89	1.00	0.98	1.00	0.48	0.41
Минимум	3.22	-0.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Стандартное отклонение	1.90	0.23	0.19	0.12	0.16	0.16	0.16	0.08	0.04
LIVE = 1	10.719	0.281	0.290	0.073	0.136	0.117	0.202	0.063	0.034
LIVE = 0	10.533	0.174	0.267	0.024	0.073	0.139	0.285	0.049	0.056

Предварительный анализ средних значений параметров банков на апрель 1998 г. указывает на некоторые различия между выжившими (LIVE = 1) и обанкротившимися (LIVE = 0) во время кризиса банками.

Наибольшие различия между средними значениями индикаторов по двум группам банков относятся к индикаторам SK/VB, GDO/VB, LA/VB, которые значительно выше для выживших банков, и к индикаторам OS/VB, RES/VB, которые значительно выше для обанкротившихся банков.

Эти различия ожидаемы для большинства индикаторов. Отношение SK/VB подобно по смыслу нормативу Н1 достаточности капитала (ЦБ РФ), высокий уровень ликвидности LA/VB и низкая доля неработающих активов OS/VB характеризуют способность банка быстро мобилизовать ресурсы. Большие резервы RES/VB указывают на проведение банком рискованной политики кредитования.

Во время кризиса государство отказалось отвечать по своим обязательствам по ГКО. Однако GDO/VB – доля государственных ценных бумаг в активах – значительно выше у устоявших банков. Можно предложить два аргумента для объяснения этого эффекта. Первый – в стабильной развитой экономике банки должны участвовать на рынке государственных облигаций. Таким образом, наличие у банка государственных ценных бумаг в активах может говорить о зрелости его менеджмента. Второй – после кризиса государство предоставляло поддержку некоторым банкам с большой долей инвестиций в ГКО, обеспечив тем самым их выживание.

Отметим, что данные в табл. 3 не указывают на существенное различие в средних значениях размера банка LNVB.

Корреляции финансовых индикаторов банков приведены в Приложении А. Размер банка отрицательно коррелирован с RES/VB, SK/VB, LA/VB. Большой банк имеет более надежных партнеров и не должен создавать большие резервы под выданные кредиты. Такой банк потенциально более устойчив и может себе позволить иметь низкие значения достаточности капитала и доли ликвидных активов. Имея много партнеров, такой банк получает более свободный доступ к рынку межбанковских кредитов и к другим источникам ресурсов.

Совокупный собственный капитал банковской системы в первом квартале 1998 г. был весьма мал по сравнению с ВВП, к тому же существовало слишком много мелких банков. В то время наиболее популярными банковскими стратегиями были сильные связи с промышленными группами или спекуляции на высокодоходном рынке ГКО.

2.2. Модели и кластеры. Для моделирования вероятности дефолта мы используем модель бинарного выбора

$$P(\text{LIVE}_i = 1) = F(x_i', \beta), \quad (1)$$

где x_i – вектор параметров (показателей) банка i . Оказалось, что logit-модель, соответствующая спецификации $F(z) = (1 + e^{-z})^{-1}$, лучше подходит к данным, чем probit-модель.

Уравнение (1) можно оценивать по всему набору банков. Однако возможна ситуация, когда при оценивании этого уравнения отдельно по двум кластерам мы получим значимый, но с разным знаком вклад некоторого параметра, в то время как оценивая уравнение по объединению кластеров, мы приходим к выводу о незначимости этого параметра. Это соображение является основным объяснением идеи рассмотрения разных моделей в кластерах банков.

Такие стандартные процедуры кластер-анализа, как наиболее часто используемая процедура k -средних, которые могут давать интересные результаты (Бобышев, 2001), не подходят для наших целей, поскольку они имеют тенденцию группировать точки наблюдений с похожими зна-

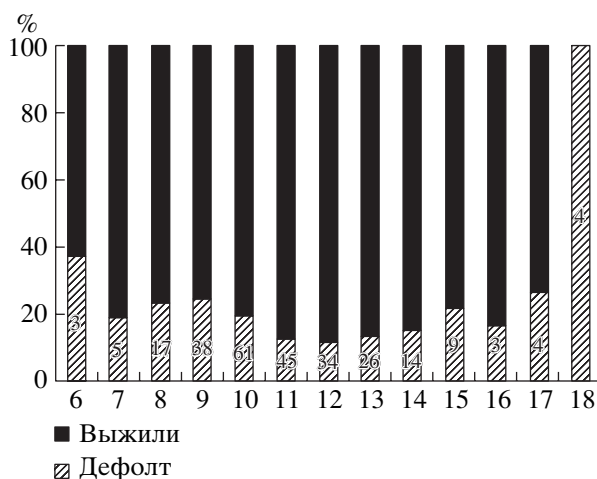


Рис. 1. Распределение дефолтов банков по размеру, LNVB.

чениями параметров в один компактный кластер. Этот алгоритм минимизирует сумму расстояний от точек до центров кластеров. Нашей же целью является построение в кластере logit-модели с наилучшим качеством подгонки.

Единственный известный нам пример использования кластер-анализа в моделировании надежности финансовых институтов – работа (Wescott, 1984), в которой автор применяет алгоритм метода k -средних в задаче прогноза рейтинга муниципальных облигаций в США в 1977 г. Он приходит к выводу, что построение отдельных моделей в каждом кластере не улучшает качество прогноза. Это говорит в пользу аргументов, приведенных ранее.

Альтернативой могла бы стать классификация банков по профилю их активности (“карманные” банки; банки, аффилированные с промышленными группами; банки, ориентированные на обслуживание экспортно-импортных

операций, и т.п.). К сожалению, такая информация отсутствует по большинству банков из выборки.

Ниже мы используем “экспертный подход”, при котором эксперт определяет параметр и два пороговых значения этого параметра. Банки таким образом разбиваются на три кластера – с малыми, средними и большими значениями данного параметра (например, размера банка). Далее logit-модель оценивается для каждого из трех кластеров. Преимуществом подхода является прозрачность интерпретации кластеров. Недостатком – субъективность в выборе пороговых значений.

Кроме того, ниже предложена автоматическая процедура кластеризации, которая выбирает кластеры, принимая во внимание качество подгонки logit-модели в каждом из кластеров. Алгоритмы, подобные предложенному, сочетающие в себе кластеризацию с одновременной подгонкой моделей, хорошо себя зарекомендовали (для совершенно другой задачи) в работах (Borodovsky, Peresetsky, 1994; Mathe et al., 1999).

2.3. Экспертный подход. Для экспертной кластеризации использовались четыре различных финансовых показателя банка. Для каждого показателя экспертно были выбраны два пороговых значения. Таким образом, получились четыре разбиения на кластеры. Каждое разбиение состоит из трех кластеров с малыми, средними и большими значениями параметра.

Суммарные активы (VB). Кластер малых банков $VB = 1\%$ состоит из банков с общими суммарными активами, составляющими 1% суммарных активов всей банковской системы. Кластер больших банков $VB = 90\%$ состоит из крупнейших банков, составляющих 90% суммарных активов всей банковской системы.

Доля государственных ценных бумаг (GDO/VB). Кластер банков, практически не участвующих на рынке ГКО, $GDO/VB < 0.01\%$, и кластер банков, активно инвестирующих в ГКО, $GDO/VB > 10\%$.

Доля кредитов нефинансовым организациям (KE/VB). Кластер “пассивных” банков – $KE/VB < 15\%$ и кластер “активных” банков – $KE/VB > 40\%$.

Капитализация (SK/VB). Кластер банков с малыми значениями отношения капитал/активы – $SK/VB < 11\%$ и кластер высококапитализированных банков – $SK/VB > 30\%$.

Распределение банков по кластерам и пересечения кластеров приведены в табл. 4.

В Приложении С приведены средние значения финансовых показателей банков по кластерам. Доля дефолтов выше среди банков с малыми значениями капитализации, долей государственных ценных бумаг и кредитов нефинансовым организациям. Дефолты в банках среднего размера встречаются реже, чем в малых и крупных банках.

Для более детального анализа разделим весь диапазон изменения размера банка (логарифма суммарных активов – LNVB) на интервалы и построим на графике доли выживших и обанкротившихся банков в каждом интервале (рис. 1). Цифры в столбиках означают число обанкротившихся банков. U-образное распределение показывает, что у банков среднего размера было больше шансов выжить во время кризиса.

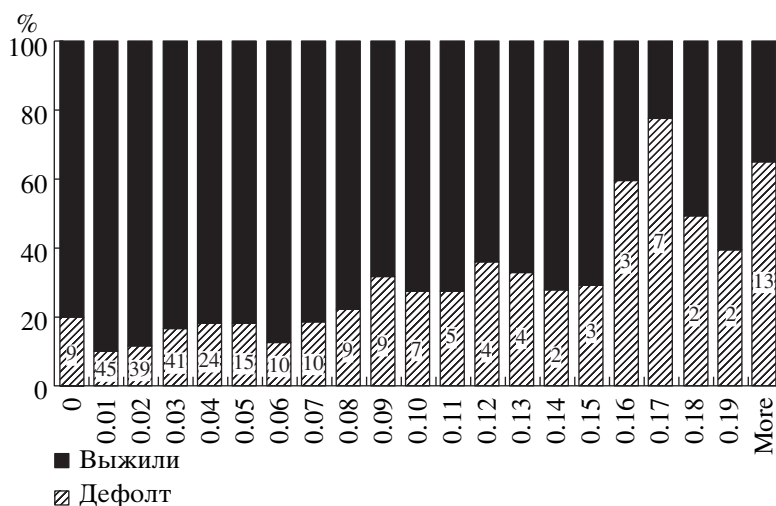


Рис. 2. Распределение дефолтов банков по доле резервов в активах, RES/VB.

Аналогичный график распределения дефолтов по доле резервов RES/VB приведен на рис. 2. На этот раз распределение монотонно возрастает, т.е. банки с высокими значениями резервов, а следовательно, с рискованной кредитной политикой, имели больше шансов потерпеть дефолт.

Распределение дефолтов по доле депозитов физических лиц VDFL/VB имеет U-образное распределение, подобное распределению на рис. 1, а распределения доли дефолтов по доле основного капитала и неработающих активов, а также по доле негосударственных ценных бумаг монотонны и похожи на распределение на рис. 2. Эти графики говорят в пользу первоначальной гипотезы, согласно которой высокие доли резервов, негосударственных ценных бумаг, неликвидных активов повышают вероятность дефолта банка во время кризиса. У банков со средними значениями доли депозитов физических лиц больше шансов выжить во время кризиса. На первый взгляд это может показаться странным, но напомним, что во время кризиса 1998 г. банки просто заморозили счета физических лиц и использовали эти деньги для выполнения своих других обязательств. Этот аргумент, по крайней мере, частично объясняет, почему более высокая доля депозитов физических лиц может соответствовать более высокой вероятности выживания во время кризиса.

В каждом из перечисленных выше кластеров была построена модель вероятности дефолта. При выборе модели использовались значения статистики качества подгонки модели LR (отно-

Таблица 4. Пересечения кластеров

Кластеры	VB = 1%	VB = 90%	GDO/VB < 0.01%	GDO/VB > 10%	KE/VB < 15%	KE/VB > 40%	SK/VB < 11%	SK/VB > 30%
VB = 1%	624	0	403	100	197	179	94	359
VB = 90%	0	261	22	93	50	54	78	34
GDO/VB < 0.01%	403	22	624	0	187	203	126	301
GDO/VB > 10%	100	93	0	378	110	51	40	149
KE/VB < 15%	197	50	187	110	392	0	89	187
KE/VB > 40%	179	54	203	51	0	425	50	194
SK/VB < 11%	94	78	126	40	89	50	268	0
SK/VB > 30%	359	34	301	149	187	194	0	615

Таблица 5. Модели для мелких и крупных банков³

Кластер мелких банков VB = 1%			Кластер крупных банков VB = 90%		
Переменная	Коэффициент	$s_x\beta_x$	Переменная	Коэффициент	$s_x\beta_x$
C	-2.05*	—	C	7.24***	—
SK/VB	1.65***	0.43	SK/VB	5.71***	0.81
LNVB	0.26**	0.27	LNVB	-0.39***	-0.46
NDO/VB	-1.67**	-0.31	NDO/VB	-5.94***	-0.59
OS/VB	-1.82**	-0.32	OS/VB	-2.66**	-0.41
KE/VB	4.50***	0.95			
(KE/VB) ²	-5.64***	-0.87			
GDO/VB	8.41***	1.02			
LA/VB	3.35***	0.65			
RES/VB	-5.09**	-0.27			
VDFL/VB	3.93**	0.29			
МакФадден R^2		0.22	МакФадден R^2		0.17
Наблюдения с LIVE = 0		132	Наблюдения с LIVE = 0		47
Наблюдения с LIVE = 1		492	Наблюдения с LIVE = 1		214
Всего наблюдений		624	Всего наблюдений		261

шение правдоподобия), МакФадден R^2 , а также z-статистики значимости коэффициентов. Некоторые из этих моделей приведены ниже.

Мелкие и крупные банки (VB = 1% и VB = 90%). Кластер мелких банков отличается высокими значениями капитализации (34%, все – 28%, крупные – 18%) и ликвидности (17%, все – 14%). Средние значения по этому кластеру не отличаются значимо от средних значений по выборке. В кластере мелких банков процент дефолтов равен 21% (все банки вместе – 17%, крупные – 18%).

Лучшие модели для двух кластеров приведены в табл. 5. Значения $s_x\beta_x$ позволяют сравнивать степень влияния переменных на вероятность дефолта (здесь β_x – оценка коэффициента, а s_x – стандартное отклонение значений переменной в кластере).

Знаки коэффициентов ожидаемые. Высокие значения капитализации и доли ликвидных активов (SK/VB, LA/VB) повышают вероятность выживания банков, в то время как высокие значения доли резервов, неликвидных активов и негосударственных ценных бумаг увеличивают вероятность их дефолта. Размер банка, LNVB, оказывает разное влияние в разных кластерах. В кластере мелких банков увеличение размера повышает вероятность выживания, в то время как для крупных банков влияние этого параметра противоположное. В самом деле, 4 крупнейших коммерческих банка (СБС-Агро, Инкомбанк, Менатеп, Российский Кредит) обанкротились в результате кризиса. Среди 28 крупнейших российских банков было 11 дефолтов.

Модель для мелких банков включает большее число параметров, значения коэффициентов при них допускают естественную интерпретацию. Разные знаки при KE/VB и (KE/VB)² указывают на существование оптимального значения доли кредитов нефинансовым организациям, равное 40%, что близко значению этого показателя по банковскому сектору в 2004 г. Среднее значение по кластеру равно 28%, это может означать, что поведение банков до кризиса не было оптимальным.

Доля депозитов физических лиц увеличивает вероятность выживания. В то время только 50 обанкротившихся банков (19%) и 320 выживших (25%) имели ненулевые объемы депозитов физических лиц.

Наиболее “влиятельными” переменными в кластере мелких банков являются GDO/VB и KE/VB. Это до некоторой степени объясняется большим числом “карманных” банков в этом кластере. В кластере крупных банков наиболее влиятельным параметром является капитализация (т.е. параметр, близкий по смыслу к Н1 – нормативу достаточности капитала, который использует ЦБ РФ).

³ Во всех таблицах *, ** и *** обозначают значимость на уровнях 10, 5 и 1%.

Таблица 6. Модели для банков с низким и высоким уровнем инвестиций в ГКО

Кластер GDO/VB < 0.01%			Кластер GDO/VB > 10%		
Переменная	Коэффициент	$s_x \beta_x$	Переменная	Коэффициент	$s_x \beta_x$
Const	-2.81***		Const	8.69***	
SK/VB	1.38***	0.38	SK/VB	6.45***	1.18
RES/VB	-3.37*	-0.19	RES/VB	-15.48**	-0.45
LNVB	0.24***	0.37	LNVB	-0.39**	-0.70
LA/VB	4.69***	0.95	OS/VB	-7.93***	-0.72
KE/VB	4.28***	0.92	NDO/VB	-5.74**	
(KE/VB) ²	-3.70*	-0.59			
МакФадден R^2		0.135	МакФадден R^2		0.22
Наблюдения с LIVE = 0		161	Наблюдения с LIVE = 0		16
Наблюдения с LIVE = 1		463	Наблюдения с LIVE = 1		362
Всего наблюдений		624	Всего наблюдений		378

Кластеры банков с малыми и большими инвестициями в государственные ценные бумаги (GDO/VB < 0.01% и GDO/VB > 10%). Средний размер (LNVB) банка с низким уровнем инвестиций в ГКО равен 9.62, в то время как для кластера банков с высоким уровнем инвестиций это значение равно 11.36 (см. Приложение С), что существенно отличается от общего среднего 10.72. Средние значения других параметров по кластерам практически не отличаются от средних по всем банкам.

Доли дефолтов в кластерах равны соответственно 26 и 4%, они существенно отличаются от общей доли дефолтов, равной 17%. Как отмечалось ранее, высокий уровень инвестиций в ГКО может означать зрелость менеджмента банка, кроме того, некоторые из таких банков получили государственную поддержку после кризиса, что может объяснять низкую вероятность дефолта в этом кластере.

Модели для двух кластеров приведены в табл. 6. Кредиты экономике, KE/VB, значимы для банков с низким уровнем инвестиций в ГКО. Оптимальное значение этого параметра составляет 58%, что гораздо выше 30% – среднего по кластеру. Возможно, этим банкам следовало бы увеличить инвестиции в экономику. Для этого кластера наиболее “влиятельными” параметрами являются ликвидные активы и кредиты нефинансовым организациям.

Для второго кластера (с высоким уровнем инвестиций в ГКО) кредиты нефинансовым организациям, KE/VB, незначимы (табл. 6). Ликвидные активы и негосударственные ценные бумаги значимы, и рост их доли повышает вероятность дефолта. Наиболее “влиятельным” параметром оказывается капитализация – SK/VB. Размер банка оказывает негативное влияние на вероятность дефолта, поскольку большинство крупнейших банков принадлежат к этому кластеру.

В Приложении D показаны результаты оценки моделей с одним набором переменных по всем кластерам. Индикаторы SK/VB, LA/VB и GDO/VB значимы для большинства кластеров.

Доля просроченных кредитов в активах, которая была указана значимой в работе (Bovenzi et al., 1983), оказалась незначимой ни в одном из кластеров. Возможным объяснением этому является тот факт, что в 1998 г. многие российские банки скрывали истинные значения этого параметра в отчетности. Российская система отчетности позволяла виртуально закрывать и снова открывать такие кредиты, что снижало значение PZS/VB в балансе банка.

Значение критерия МакФадден R^2 показывает, что в некоторых кластерах качество подгонки выше, чем для всей выборки в целом. Однако является ли это улучшение статистического критерия качества подгонки значимым для прогнозной силы модели?

Прогноз внутри выборки (in-sample). Когда оценки коэффициентов модели получены, можно рассчитать прогнозные вероятности выживания $\hat{p}_i$ для каждого банка i из выборки. Для того чтобы получить прогноз, необходимо выбрать пороговое значение. Если $\hat{p}_i < c$, то прогнозируется дефолт банка, если же $\hat{p}_i \geq c$, то прогнозируется, что банк выживет. Назовем ошибкой I рода ситуацию, когда банк с прогнозом выживания терпит дефолт. Ошибка II рода состоит в том,

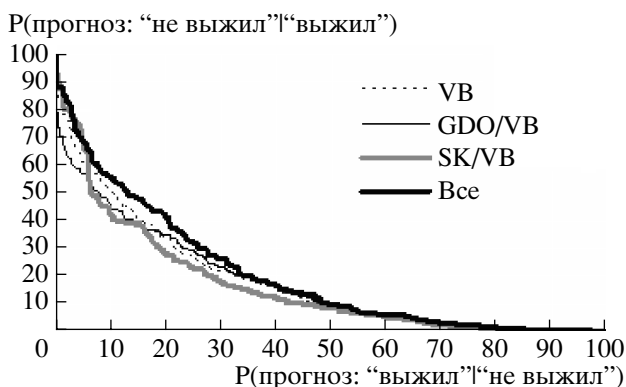


Рис. 3. Вероятности ошибок I–II рода.

лученным при кластеризации по значениям показателей GDO/VB, VB и SK/VB. В наиболее интересной с практической точки зрения области малых вероятностей ошибок I рода заметное улучшение прогноза достигается при кластеризации по доле государственных ценных бумаг в валюте баланса, GDO/VB. Конечно, надо иметь в виду, что модель с кластеризацией более гибкая, поскольку содержит в 3 раза больше коэффициентов. Графики, построенные по отдельным кластерам, выглядят более убедительно, так, например, в кластере $GKO > 10\%$ в области вероятности ошибок I рода, меньших 20%, модель с кластеризацией уменьшает вероятность ошибки II рода на 15–35%.

Прогноз вне выборки (out-of-sample). Построение модели по части выборки и затем ее тестирование на оставшейся части выборки является более надежной оценкой качества модели, чем ее тестирование на той же выборке, на которой она была построена. Для этого используется следующая процедура. При помощи датчика случайных чисел выборка разбивается на две части: основную группу (1465 банков) и контрольную (100 банков). Модель оценивается на данных основной группы, а затем доли ошибок I–II рода рассчитываются по контрольной группе. Усредненные по 1000 таких разбиений результаты приведены на рис. 4. Графики на рис. 4 практически не отличаются от графиков на рис. 3.

Для демонстрации способности модели предсказывать “надежные” и “ненадежные” банки рассчитаем пропорции n_1 и n_2 банков, обанкротившихся реально, среди 10 “самых ненадежных” и “самых надежных” с точки зрения модели, банков в каждой контрольной группе. При идеальном прогнозе $n_1 = 100\%$ и $n_2 = 0\%$. Результаты по каждой из 17 серий по 1000 попыток в каждой усреднились, и результаты приведены на рис. 5.

Наилучшие результаты прогноза “надежных” банков были получены при кластеризации по доле государственных ценных бумаг GDO/VB – в среднем менее 1% ошибок среди 10 лучших банков в контрольной группе. Для выявления проблемных банков несколько лучше работают модель без кластеризации и модель с кластеризацией по капитализации SK/VB. Эти модели дают примерно 57% корректно предсказанных дефолтов среди 10 “самых ненадежных” банков в контрольной группе.

2.4. Автоматическая классификация. Модель вероятности дефолта, которая включает две разные logit-модели в двух кластерах, состоит в следующем.

Рассмотрим logit-модель, которая разделяет банки на два кластера. Пусть вероятность того, что банк с набором параметров z принадлежит к первому кластеру, равна $F(z'\gamma)$, где γ – вектор коэффициентов, а F – логистическая функция распределения. Вероятность того, что банк принадлежит второму кластеру, равна, соответственно, $1 - F(z'\gamma)$.

Таблица 7. Модель автоматической классификации

Переменная	Коэффициенты γ
Const	–3.14
SK/CA	13.29***
LnVB	–0.0129***

что банк с прогнозом дефолт выживает. Очевидно, потери от ошибки I рода значительно больше, чем от ошибки II рода. Выбор порогового значения определяется балансом стоимостей ошибок I и II рода для инвестора (см. обсуждение компромисса между ошибками I и II рода в работе (Bovenzi et al., 1983)). Для каждого значения порога c получается пара $(p_{i1}(c), p_{i2}(c))$ – вероятности (доли) ошибок I, II рода по выборке. Эти точки образуют график. Одна модель равномерно лучше другой, если ее график лежит ниже графика второй модели.

На рис. 3 приведены графики вероятностей ошибок I–II рода, полученные по модели, построенной по всей выборке, а также по моделям, построенным отдельно по кластерам, по-

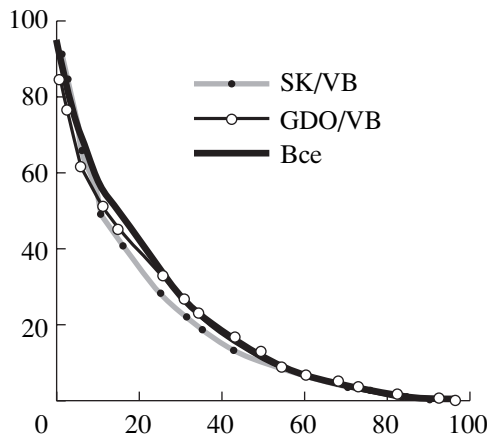


Рис. 4. Ошибки I–II рода.

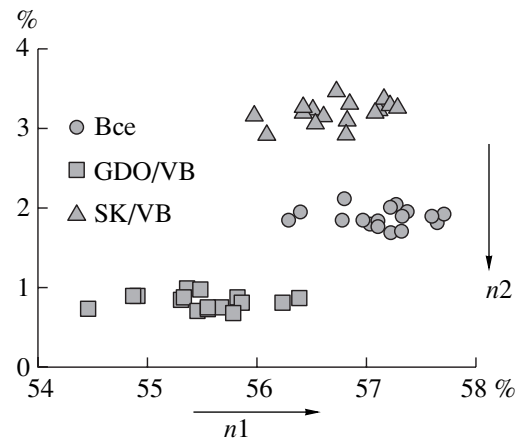


Рис. 5. Доли обанкротившихся банков.

В каждом из двух кластеров построим модель вероятности выживания. Пусть $F(x'\beta_1)F(x'\beta_2)$ – вероятность банка с параметрами x выжить при условии, что он принадлежит первому (второму) кластеру. Различные наборы параметров могут использоваться в модели классификации и моделях выживания.

Вклад банка j в функцию правдоподобия равен

$$L_j = F(z'\gamma)(F(x'\beta_1)^{y_j}(1 - F(x'\beta_1))^{1-y_j}) + (1 - F(z'\gamma))(F(x'\beta_2)^{y_j}(1 - F(x'\beta_2))^{1-y_j}), \quad (2)$$

где $y_j = 1$, если банк выжил, и 0, если он обанкротился. Оценки параметров модели β_1 , β_2 и γ получаются максимизацией логарифмической функции правдоподобия

$$\ln L(\gamma, \beta_1, \beta_2) = \sum_j \ln L_j \rightarrow \max_{\gamma, \beta_1, \beta_2}. \quad (3)$$

Поскольку нет гарантии, что функция (3) имеет глобальный максимум, возникает естественная проблема неединственности оценок. Например, если набор $(\gamma, \beta_1, \beta_2)$ является решением, то и набор $(-\gamma, \beta_2, \beta_1)$ также является решением.

После отбора моделей мы пришли к выводу, что наилучшей является модель, в которой векторы параметров равны $z = \{SK/CA, LNVB\}$ и $x = \{SK/VB, RES/VB, KE/VB, GDO/VB, LA/VB\}$. Оценка модели разбиения на кластеры представлена в табл. 7. Отметим, что полученный алгоритм кластеризации похож на разделение банков на две группы в соответствии с инструкцией № 1 ЦБ РФ, в которой более низкий порог достаточности капитала устанавливается для больших банков. Полученные результаты автоматической классификации могут быть использованы ЦБ РФ для установления более гибких требований к достаточности капитала.

В качестве порога кластеризации используется 0.5, т.е. банк принадлежит к первому кластеру, если $F(z'\gamma) < 0.5$, и ко второму – в противном случае. В каждом кластере logit-модель оценивается отдельно. Графики вероятностей ошибок I–II рода для модели с набором параметров x без кластеризации и для модели, основанной на кластеризации (2)–(3), представлены на рис. 6. На рисунке видно, что использование автоматической кластеризации повышает прогнозную силу модели.

Модель (2)–(3) может быть легко модифицирована для автоматической классификации на несколько кластеров.

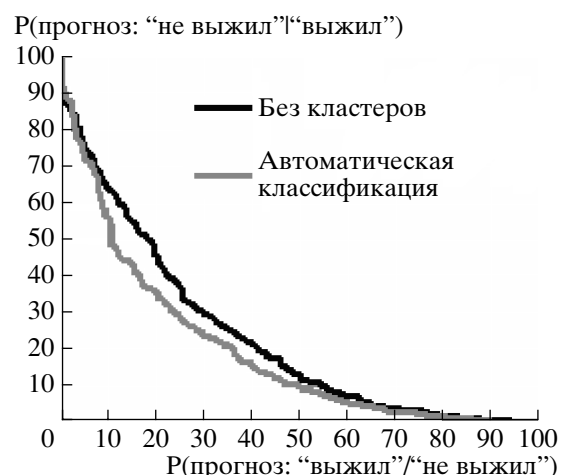


Рис. 6. Автоматическая классификация.

Таблица 8. Число дефолтов среди n “надежных” банков

n	100	200	300	400	500	600	700	800
Среднее	17	34	50	67	84	101	117	134
Базисная модель	0	5	9	12	18	26	37	53
GDO/VB	0	0	1	4	8	16	24	33
VB	1	2	6	8	13	22	29	39
SK/VB	1	3	8	13	16	16	21	28
KE/VB	0	6	9	15	18	25	33	43
Автоматическая классификация	1	3	6	11	18	23	28	34

Таблица 9. Число дефолтов среди n “ненадежных” банков (всего 263 дефолта)

n	50	100	150	200	250	300	350
Среднее	8	17	25	34	42	50	59
Базисная модель	39	67	91	115	128	138	152
GDO/VB	41	71	92	116	129	142	152
KE/VB	42	73	94	114	129	144	158
SK/VB	41	74	96	116	136	153	166
VB	42	68	92	114	128	144	155
Автоматическая классификация	39	75	101	119	130	146	156

В табл. 8 и 9 представлено сравнение прогнозной силы различных моделей с кластеризацией для прогноза внутри выборки (in-sample). Как было показано выше, прогнозы вне выборки и внутри выборки практически совпадают по оценке прогнозной силы моделей.

В табл. 8 показано сравнение моделей с точки зрения инвестора. Для каждой модели выбрано 100 (200 и т.д.) “самых надежных” банков и среди них подсчитано число реально обанкротившихся.

Например, столбец 3 (с заголовком 200) содержит информацию по 200 “самым надежным”, с точки зрения каждой из моделей, банкам. Строка “среднее” показывает ожидаемое число дефолтов в случайной выборке банков. Строка “базисная” представляет результат модели m_0 из таблицы Приложения D; только 5 банков обанкротились. В строках, помеченных GDO/VB, VB, SK/VB и KE/VB, представлены результаты моделей с тем же набором регрессоров, оцененных отдельно по каждому из кластеров, полученных путем “экспертной” кластеризации по соответствующим параметрам.

В последней строке представлены результаты описанной выше процедуры с автоматической классификацией. Из данных таблицы становится очевидно, что использование моделей существенно снижает вероятность выбора в качестве партнера такого банка, который обанкротился в будущем. Кластеризация по доле государственных ценных бумаг сводит число обанкротившихся банков среди 200 “самых надежных” к 0. При сравнении результатов надо иметь в виду, что модель с автоматической классификацией (на два кластера) содержит 15 параметров, базисная модель – 6 параметров, а модель с экспертной классификацией (на 3 кластера) – 18 параметров.

В табл. 9 приведено сравнение прогнозной силы моделей с точки зрения органов банковского надзора. Для каждой модели было выбрано 100 (200 и т.д.) “самых ненадежных”, с точки зрения модели, банков, и показано число дефолтов среди них. Например, выделив 250 “самых ненадежных” банков (16% их общего числа), мы обнаруживаем 136 дефолтов (52% общего числа дефолтов). Это позволяет сузить круг потенциально проблемных банков и, возможно, сэкономить органам банковского надзора значительное время и средства для проведения инспекций на месте.

3. МАКРОПЕРЕМЕННЫЕ

В работах, посвященных моделям вероятности дефолта банков (см., например (Martin, 1977; Estrella et al., 2000; Kolari et al., 2002)), в моделях не используются переменные макроокружения. Во всех этих работах рассматривается банковская система США, которая функционирует в усло-

виях относительно стабильной экономики. Поскольку экономическая ситуация в 1996–2003 гг. в России была далека от стабильной, представляется, что включение в модели макроиндикаторов может повысить прогнозную силу моделей.

Существует лишь несколько работ, где в моделях вероятности дефолта фирм, кредитов или облигаций используются макроиндикаторы. В работе (Engelmann, Porath, 2003) показано, что использование в модели параметров прироста реального ВВП и прироста денежной массы (M3) улучшает качество logit-модели вероятности дефолта германских компаний в 1989–2000 гг. Уровень безработицы используется в моделях вероятности дефолта по ипотечным кредитам в США в (Lawrence, Smith, 1992). В работах (Головань и др., 2004; Peresetsky et al., 2004) макроиндикаторы используются в моделях вероятности дефолта российских банков.

Многие публикации Базельского комитета подчеркивают роль макроэкономического окружения в оценке риска дефолта (см., например (Amato, Furfine, 2003; Borio, 2003; Segoviano, Lowe, 2002)).

В работе (Godlewski, 2004) рассматриваются переходные экономики (не включая Россию). Автор указывает на то, что использование макроиндикаторов может улучшить качество скоринговых моделей банков.

Макроиндикаторы используются в работах, изучающих причины банковских и финансовых кризисов. В работе (Demirguç-Kunt, Detragiache, 1998) объединенная logit-модель была применена для изучения банковских кризисов в развивающихся и индустриальных странах в 1980–1994 гг. Авторы показали, что рост ВВП, реальные процентные ставки, инфляция, условия торговли значимы во всех спецификациях моделей. Они не обнаружили независимый эффект от обменного курса, по-видимому, инфляция и условия торговли уже покрывают этот эффект.

Панельная probit-модель используется в работе (Komulainen, Lukkarila, 2003) для изучения причин финансовых кризисов в 31 стране с переходной экономикой в 1980–2001 гг. Авторами показано, что многие макроиндикаторы важны для прогноза финансового кризиса, в том числе некоторые, не указанные в (Demirguç-Kunt, Detragiache, 1998), например, уровень безработицы, объем промышленной продукции. Поскольку финансовые и банковские кризисы часто связаны, все эти макроиндикаторы также могут быть важны для моделей вероятности дефолта банка.

3.1. Данные. Финансовые показатели банков. В этом разделе используются финансовые показатели квартальных отчетов российских банков за период 1996–2002 гг.⁴ Переменная LIVE сконструирована на основании определения дефолта банка, сформулированного в п. 2.1.

Для того чтобы избежать проблемы коррелированности наблюдений и увеличить долю дефолтов, мы сокращаем выборку. Цель прореживания выборки состоит в том, чтобы сохранить всю информацию о дефолтах и чтобы наблюдения для каждого банка отстояли по времени друг от друга на два года.

Для каждого обанкротившегося банка мы берем момент банкротства t (время измеряется в кварталах) и полагаем LIVE = 0. Это наблюдение привязывается к значениям финансовых показателей банка (табл. 1) и макроиндикаторов (табл. 10) в момент $t - 8$ ⁵. Далее берется тот же банк в момент $t - 8$ и полагается LIVE = 1. К этому наблюдению привязываются финансовые показатели банков и макроиндикаторы на момент $t - 16$. Движение назад по времени с шагом в два года продолжается, пока в выборке есть данные об этом банке.

Для выживших банков процедура несколько иная. Произвольно выбираем квартал t из 8 кварталов за период 2001–2002 гг., полагаем LIVE = 1 и связываем с этим наблюдением финансовые показатели банков и макроиндикаторы на момент $t - 8$. Движение назад по времени с шагом в два года продолжается так же, как и для обанкротившихся банков.

Подобная процедура прореживания выборки приводит, вообще говоря, к смещенным оценкам параметров (Scott, Wild, 1986). Однако смещенной оказывается оценка только константы, оценки же коэффициентов при переменных несмещенные, поэтому это не меняет экономической интерпретации результатов регрессии. К тому же смещение оценки свободного члена не влияет на оценку прогнозной силы модели, поскольку при сравнении моделей рассматриваются все возможные значения порога c .

В результате описанной выше процедуры прореживания получается выборка, состоящая из 3158 наблюдений, среди которых есть 255 дефолтов (8.07%). Распределение дефолтов приведено

⁴ Данные получены от Информационного агентства “Мобиле”.

⁵ Как и в п. 2.1, использование двухгодичного лага между финансовыми показателями банка и его статусом дает наилучший результат.

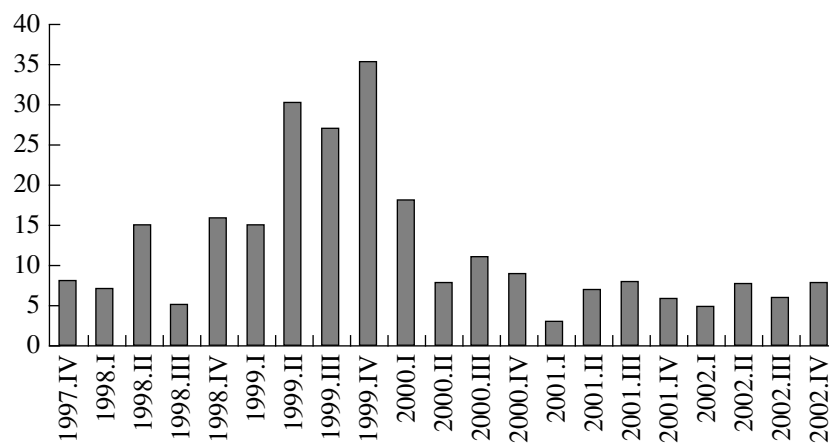


Рис. 7. Распределение дефолтов банков в прореженной выборке, разд. 3.

на рис. 7. Это распределение не совпадает с общим распределением дефолтов (Приложение В), поскольку в выборку включены только те банки, у которых имеются данные об их финансовых показателях за два года до дефолта.

Макроэкономические индикаторы. Список макроиндикаторов, рассмотренных для включения в модели, представлен в табл. 10. Выбор макроиндикаторов определен мнением экспертов и

Таблица 10. Макроиндикаторы

Индекс реального ВВП (с поправкой на сезонность)*, %	VVP
Индекс потребительских цен, %	CPI
Дефлятор**	Defl
Число безработных, млн. чел.	UNEMPLN
Уровень безработицы, %	UNEMPLP
Индекс реальных инвестиций в основной капитал (с поправкой на сезонность)***, %	INV
Обменный курс, руб./долл.	ERATE
Отношение объема экспорта к импорту	EXP/IMP
Прирост объема промышленной продукции к соответствующему периоду предыдущего года, %	PRPROD
Прирост реальных денежных доходов населения, %	REALINC
Приращение обменного курса ERATE за год	DERATE
Приращение обменного курса ERATE за квартал	DERATE1
Изменение темпов роста показателя ВВП за год, %	DVVP
Изменение темпов роста показателя ВВП за квартал, %	DVVP1

Примечание: * 1994.01 = 100; ** на базе CPI; *** 1993. I = 100.

Таблица 11. Статистика финансовых индикаторов банков

Значения	KE/VB	NDO/VB	PNA/VB	CFB/VB	$\ln(VB/Defl)$	$\ln(VB/Defl)^2$	SK/VB	BP/VB
Среднее	0.300	0.087	0.105	0.030	10.37	111.3	0.301	0.016
Максимум	0.988	0.978	0.943	0.933	16.98	288.2	0.997	0.783
Минимум	0.000	0.000	0.000	0.000	3.98	15.9	-0.729	-0.716
Стандартное отклонение	0.205	0.143	0.121	0.084	1.95	41.7	0.221	0.057

Таблица 12. Корреляция макроиндикаторов

Значения	CPI	ERATE	EXP/IMP	REALINC	PRPROD	UNEMPLN	UNEMPLP	VVP
CPI	1	-0.040	-0.010	-0.521	-0.586	0.287	0.301	-0.441
ERATE	-0.040	1	0.933	0.043	0.609	0.037	0.081	0.680
EXP/IMP	-0.010	0.933	1	0.156	0.534	-0.069	-0.028	0.669
REALINC	-0.521	0.043	0.156	1	0.401	-0.816	-0.822	0.711
PRPROD	-0.586	0.609	0.534	0.401	1	-0.178	-0.153	0.723
UNEMPLN	0.287	0.037	-0.070	-0.816	-0.178	1	0.996	-0.618
UNEMPLP	0.301	0.081	-0.028	-0.822	-0.153	0.996	1	-0.584
VVP	-0.441	0.680	0.669	0.711	0.723	-0.618	-0.584	1

Таблица 13. Модели с макроиндикаторами

Переменная	Коэффициент		
	базисная модель	макромодель 1	макромодель 2
Const	0.150	-0.847	-0.331
BP/VB	1.226**	1.541**	1.663***
KE/VB	-1.188***	-0.976***	-0.930***
NDO/VB	-1.008***	-1.247***	-1.394***
PNA/VB	-1.346***	-1.223***	-1.204***
CFB/VB	-0.546	-0.277	-0.113
ln(VB/Defl)	0.376***	0.367***	0.368***
ln(VB/Defl) ²	-0.0181***	-0.0181***	-0.0184***
EXP/IMP	—	0.621***	—
ERATE	—	—	0.0346***
Логарифм правдоподобия	-845.11	-807.73	-799.74
LR-статистика (8, 9 степеней свободы)	81.99	156.76	172.72
Критерий Акаике	0.5403	0.517	0.512
МакФадден R ²	0.0463	0.088	0.097

работами, в которых было обнаружено значимое влияние макроиндикаторов⁶. Для некоторых макроиндикаторов рассматривался также темп их прироста.

В табл. 11 представлены описательные статистики некоторых финансовых индикаторов банков. Все финансовые показатели измеряются в тысячах рублей. Среднее значение общих активов банка, измеренное в долларах США по историческому курсу, составляет всего около 100 млн. долл. Эта величина весьма мала не только по сравнению с иностранными банками, но и по сравнению с крупными российскими банками. Это является еще одной причиной того, чтобы измерять размер банка как логарифм активов, что позволит уменьшить разброс этого параметра в выборке. Кроме того, поскольку данные в выборке распределены во времени, представляется разумным взять в качестве размера банка дефлированное значение активов, т.е. $\ln(\text{VB}/\text{Defl})$.

Для того чтобы решить, какие макроиндикаторы могут быть включены в модель, рассмотрим их корреляции (табл. 12). Большинство макроиндикаторов высоко коррелированы, поэтому включение в модель двух или более макроиндикаторов может привести к мультиколлинеарности. В табл. 12 выделены пары макроиндикаторов с наименьшей корреляцией, которые потенциально могут быть включены в модель одновременно.

⁶ Для банковских и финансовых кризисов макроиндикаторы были взяты из работ (Demirguç-Kunt, Detragiache, 1998) и (Komulainen, Lukkarila, 2003), для дефолта фирм – (Engelmann, Porath, 2003), для банков – (Головань и др., 2004).

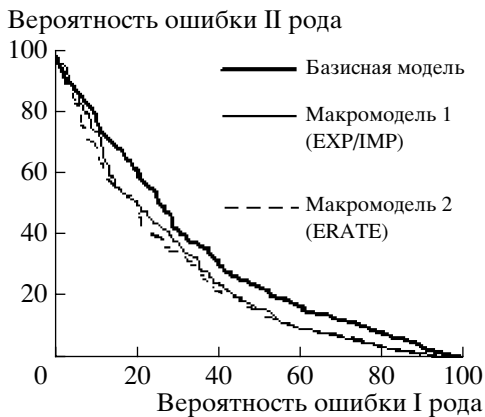


Рис. 8. Вероятности ошибок I–II рода.

Лучшие две модели, отличающиеся от базисной модели добавлением макроиндикаторов (табл. 10), представлены в столбцах 3 и 4 табл. 13. Макромодель 1 включает отношения экспорта к импорту, а макромоделю 2 включает обменный курс рубль/доллар. По всем статистическим критериям (логарифм правдоподобия, Акаике, МакФадден R^2) обе макромодели превосходят базисную. Некоторое дальнейшее улучшение может быть достигнуто путем включения в модель двух макропеременных. Однако не очевидно, что эти дальнейшие улучшения экономически значимы. Следующие два раздела посвящены изучению этого вопроса.

Знаки коэффициентов моделей в табл. 13 соответствуют экономической рациональности. Отношение дохода до вычета налогов к активам (BP/VB), которое может использоваться как мера качества менеджмента, оказывает положительное влияние на вероятность выживания банка. Доля кредитов нефинансовым организациям в активах (KE/VB) оказывает отрицательное влияние. Это отличается от вывода разд. 2. Однако следует заметить, что в отличие от разд. 2, в котором исследуется финансовый кризис, в данной модели используются данные пятилетнего периода, в котором доля кредитов значительно менялась вместе с макроокружением. Доли негосударственных ценных бумаг и неработающих активов (NDO/VB и PNA/VB) дают отрицательный эффект, что говорит о недостаточно качественном управлении активами. Модель указывает на “оптимальный” размер банка, $\ln(VB/Defl) = 10.44$, что несколько выше среднего значения этого параметра по выборке. Неожиданно, но капитализация SK/VB оказывается не значимой при включении в модель. Это объясняется мультиколлинеарностью, поскольку набор регрессоров уже включает много долей активов (очевидно, сумма всех долей равна 1).

Включение в модель макропеременных увеличивает абсолютные значения маржинального эффекта BP/VB и NDO/VB и уменьшает абсолютные значения маржинального эффекта KE/VB , PNA/VB . Маржинальный эффект размера банка практически не изменяется.

Положительное значение коэффициента при EXP/IMP можно объяснить следующим образом. Рост отношения экспорт-импорт характеризует более здоровую экономику и, следовательно, макроэкономическую стабильность. В этих обстоятельствах можно ожидать и роста стабильности банковской системы.

Менее очевидна интерпретация положительного коэффициента при обменном курсе $ERATE$. В России рост обменного курса (ослабление рубля) всегда ассоциировался с макроэкономической дестабилизацией. Однако высокий обменный курс приводит к увеличению отношения экспорт-импорт. Эти два индикатора имеют высокую (положительную) корреляцию (табл. 12). Возможно, на рассматриваемом отрезке времени доминирует второй эффект.

3.3. Сравнение моделей: ошибки I–II рода. Графики, полученные по трем моделям (табл. 13) путем расчета вероятностей ошибок первого и второго рода для каждого порога c , представлены на рис. 8. На графике видно, что включение макропеременных улучшает прогнозную силу моделей, однако поскольку графики, соответствующие двум моделям с макропеременными, имеют пересечения, ни одна из двух моделей не доминирует над другой.

Прогноз внутри выборки. Для того чтобы оценить качество прогноза внутри выборки, по каждой из моделей выбраны 100 (500) “самых надежных” и “самых ненадежных” с точки зрения модели банков. Доли реально обанкротившихся банков представлены в табл. 14.

3.2. Модели с макроиндикаторами. В данном пункте рассматривается вопрос о том, насколько макропеременные могут улучшить прогнозную силу моделей вероятности дефолта банков. Базисная модель, где макроиндикаторы не используются, выбрана, как и в разд. 2, на основе статистических критериев: z -статистики коэффициентов, МакФадден R^2 , критерия Акаике и, конечно, с учетом экономической интерпретируемости. Базисная модель представлена во втором столбце табл. 13. Знаки коэффициентов соответствуют предварительным ожиданиям.

При оценивании моделей используется объединенная (pooled) probit-модель, которая здесь дает несколько лучшие результаты, чем объединенная logit-модель. Logit-модель с фиксированным эффектом и probit-модель со случайным эффектом дают практически тот же результат. Тесты показывают приемлемость объединенной модели.

Таблица 14. Прогноз доли обанкротившихся банков внутри выборки

Банки	Размер выборки	Базисная модель, %	Макромодель 1 (EXP/IMP), %	Макромодель 2 (ERATE), %
“Самые ненадежные”	100	8.2	12.2	14.1
	500	34.5	45.5	46.3
“Самые надежные”	100	1.2	1.2	0.4
	500	7.5	6.3	6.3

Таблица 15. Прогноз доли обанкротившихся банков вне выборки

Банки	Размер выборки	Базисная модель, %	Макромодель 1 (EXP/IMP), %	Макромодель 2 (ERATE), %
“Самые ненадежные”	10	7.3	11.9	12.5
	50	34.5	46.6	47.0
“Самые надежные”	10	1.3	1.5	0.6
	50	8.1	6.9	6.3

Таблица 16. Классификация результатов прогноза

Банки	Обанкротившиеся	Необанкротившиеся
Классифицированные как “проблемные” X_c^-	I. m_c	II. $n_c - m_c$
Классифицированные как “надежные” X_c^+	III. $M - m_c$	IV. $N - M - (n_c - m_c)$

Напомним, что в случайной выборке 100 банков мы в среднем получили бы $100/3158 = 3.2\%$ общего числа 255 дефолтов, а при случайном выборе 500 банков мы получили бы в среднем 15.8% общего числа дефолтов. При использовании моделей в целях банковского надзора видно, что анализ 500 “самых ненадежных” банков по базовой модели помог бы выявить 34.5% банков, которые обанкротились бы в течение следующих двух лет. Добавление в модель макропеременной увеличивает эту долю до 46.3% . Для инвестора выбор 100 “самых надежных” по модели банков позволяет сократить число обанкротившихся в выборке банков с $355 \times 3.2\% = 8$ (при случайном выборе) до 3 (базисная модель) или до 1 (макромодель 2). Эти результаты сравнимы с результатами, полученными в разд. 2, но для моделей с меньшим числом параметров.

Прогноз вне выборки. Рассмотрим следующую процедуру оценки эффекта модели. Случайно выбранные 300 наблюдений удаляются из исходной выборки и параметры трех моделей оцениваются по оставшимся наблюдениям. Затем выбранные 300 банков упорядочиваются в соответствии с каждой из моделей и для каждого упорядочивания рассматриваются наборы 10 (50) самых “надежных” и самых “проблемных” банков из 300. Затем для каждого набора подсчитывается, какая доля реально обанкротившихся банков в этой выборке из 300 банков попала в данный набор. Процедура повторяется 1000 раз, и усредненные результаты представлены в табл. 15. Как и в разд. 2, здесь не обнаружено существенной разницы между оценкой прогнозной силы модели внутри и вне выборки.

3.4. Сравнение моделей: эвристические критерии. Как отмечалось выше, цена ошибки I рода (классифицировать “проблемный” банк как “надежный”) для инвестора значительно выше цены ошибки II рода. Если бы было известно соотношение между ценами ошибок I и II рода, то можно было бы найти оптимальный выбор порога c , который минимизирует линейную функцию потерь, определенную на графике ошибок I–II рода. В данном разделе предлагаются два эвристических критерия, основанные на двух (весьма грубых) моделях поведения инвестора.

Пусть инвестор использует модель и порог c для классификации банков. Обозначим через X_c^- множество банков, классифицированных как “проблемные” (т.е. полученная по модели про-

гнозная вероятность дефолта $\hat{p} < c$), а через X_c^+ – множество банков, классифицированных как “надежные” ($\hat{p} \geq c$). Обозначения для числа банков во всех возможных группах приведены в табл. 16.

Рассмотрим “наивного” (не использующего модель) инвестора, который инвестирует (кладет депозиты) S денег в банки. Этот инвестор может использовать “равномерную” стратегию, инвестируя равные части S/N во все банки, или “пропорциональную”, при которой размер вклада пропорционален размеру банка, т.е. вклад в банк k равен $S_k = S (VB_k/\Sigma_{all})$, где VB_k – размер банка k , а $\Sigma_{all} = \sum_{j=1}^N VB_j$. Первая стратегия моделирует инвестора, который стремится диверсифицировать свои вклады, а вторая моделирует поведение всего множества инвесторов.

Обозначим через r процентную ставку по депозитам, постоянную по времени и одинаковую для всех банков (ниже мы примем $r = 15\%$ или 20% , что примерно соответствует средним ставкам российских банков в рассматриваемый период). Предположим, что вклады в обанкротившиеся впоследствии банки полностью пропадают. При этих предположениях чистый доход при “равномерной” и “пропорциональной” стратегиях равен, соответственно:

$$\frac{S}{N}(N-M)r - \frac{S}{N}M = S \frac{r(N-M) - M}{N}, \quad (4)$$

$$\frac{S}{\Sigma_{all}}(r\Sigma_{II+IV}(c) - \Sigma_{I+III}(c)). \quad (5)$$

(Суммирование производится по группам банков, указанным в табл. 16; так, например, $\Sigma_{I+III}(c) = \sum_{I+III} VB_j$.)

Рассмотрим теперь инвестора, использующего модель. Такой инвестор, используя модель и порог c , классифицирует банки на “надежные” и “проблемные/ненадежные”. Затем он следует одной из двух стратегий (“равномерная” или “пропорциональная”), однако вкладывает деньги только в “надежные” банки, а деньги, которые “наивный” инвестор вложил бы в “проблемные” банки, просто оставляет у себя⁷.

Предположим, что инвестор, использующий модель, максимизирует функцию полезности, равную ожидаемому превышению его дохода по сравнению с доходом “наивного” инвестора. Эта функция полезности равна:

– для “равномерной” стратегии

$$PR_U(c) = \frac{m_c}{N} - r \frac{n_c - m_c}{N}, \quad (6)$$

– для “пропорциональной” стратегии

$$PR_p(c) = \frac{\Sigma_I(c)}{\Sigma_{all}} - r \frac{\Sigma_{II}(c)}{\Sigma_{all}}. \quad (7)$$

Конечно, инвестор выбирает оптимальное значение порога c и, таким образом, получает два эвристических критерия для сравнения моделей:

$$PR_U = \max_{0 < c < 1} PR_U(c), \quad PR_p = \max_{0 < c < 1} PR_p(c). \quad (8)$$

В табл. 17 представлены статистические и эвристические критерии для сравнения моделей – как для базисной модели (номер 0, ряд 1), так и для 24 моделей, отличающихся от базисной набором дополнительных регрессоров. Модели 1–14 включают в качестве регрессора одну дополнительную макропеременную, а модели 15–22 – две таких макропеременных. Для сравнения в таблицу внесена модель 23, включающая фиктивные переменные – индикаторы всех кварталов. Эта модель показывает предел улучшения модели, который может быть достигнут путем включения в нее макропеременных. Модель 24 содержит фиктивную переменную CRISIS – индикатор периода после кризиса августа 1998 г.

⁷ В том случае, если инвестор, использующий модель, заранее намерен вложить все свои деньги в банк, оптимальным решением будет вложить все деньги в один самый “надежный” банк.

Таблица 17. Оценки качества моделей

N	Модель	Статистики			Эвристические критерии			
		МакФадден R^2	Log L	Akaike	PR_U ($r = 0.15$)	PR_P ($r = 0.2$)	PR_U ($r = 0.15$)	PR_P ($r = 0.2$)
0	Базисная	0.046	-845.1	0.540	1.04	0.34	10.00	7.97
1	EXP/IMP	0.088	-807.7	0.517	2.01	1.50	13.32	12.14
2	ERATE	0.097	-799.7	0.512	2.02	1.46	12.79	11.28
3	CPI	0.055	-837.2	0.536	1.34	0.60	10.50	8.93
4	VVP	0.060	-833.0	0.533	1.10	0.39	10.44	8.61
5	INV	0.049	-842.7	0.539	1.14	0.28	10.14	8.34
6	DERATE	0.062	-830.8	0.532	1.35	0.64	10.21	7.91
7	DVVP	0.055	-837.7	0.536	1.01	0.46	10.08	8.08
8	DERATE1	0.051	-841.0	0.538	1.17	0.45	10.34	8.69
9	DVVP1	0.064	-829.4	0.531	0.99	0.57	9.75	7.81
10	DERATE/ERATE	0.056	-836.5	0.535	1.23	0.52	10.21	8.21
11	DERATE1/ERATE	0.049	-842.6	0.539	1.14	0.33	10.21	8.23
12	REAL_INC	0.048	-843.3	0.540	0.97	0.27	10.10	8.17
13	UNEMPLP	0.046	-845.0	0.541	0.94	0.22	9.99	7.96
14	UNEMPLN	0.046	-845.1	0.541	1.00	0.31	10.01	7.98
15	ERATE, CPI	0.106	-792.2	0.508	2.12	1.51	12.36	11.17
16	ERATE, VVP	0.100	-797.7	0.512	2.15	1.51	12.23	10.88
17	EXP/IMP, CPI	0.095	-801.6	0.514	2.04	1.56	13.54	12.40
18	EXP/IMP, VVP	0.089	-807.3	0.518	2.01	1.48	13.30	12.09
19	ERATE*(LNI/TA), ERATE	0.100	-797.1	0.511	1.96	1.32	12.72	11.20
20	ERATE*(NWA/TA), ERATE	0.098	-799.0	0.512	2.04	1.43	12.70	11.60
21	EXP/IMP*(LNI/TA), EXP/IMP	0.092	-805.0	0.516	1.87	1.37	13.26	12.23
22	EXP/IMP*(NWA/TA), EXP/IMP	0.089	-806.8	0.517	1.96	1.49	13.33	12.11
23	Индикаторы кварталов	0.129	-772.1	0.507	2.16	1.65	13.51	12.47
24	CRISIS	0.105	-793.1	0.508	2.14	1.50	12.62	11.12

Результаты, приведенные в табл. 17, показывают, что включение макропеременных улучшает статистические критерии и почти всегда улучшает эвристические критерии. Среди моделей 1–14 лучшими являются модели 1 и 2, включающие обменный курс ERATE и отношение экспорт-импорт EXP/IMP. В этих моделях эвристический критерий PR_U увеличивается по сравнению с базисной моделью с 1 до 2%, а критерий PR_P – с 10 до 13%.

В нескольких случаях при добавлении макропеременной значения эвристических критериев уменьшаются (например, при добавлении в модель 9 темпа роста ВВП, (DVVP1)). Дело в том, что модели оцениваются методом максимального правдоподобия, максимизирующего функцию правдоподобия, что автоматически означает максимизацию МакФадден R^2 , но не эвристических критериев.

Включение двух макропеременных (модели 15–18) или перекрестных членов (модели 19–22) лишь незначительно улучшают статистические и эвристические критерии. Значения статистических и эвристических критериев модели 23 весьма близки к их значениям для моделей 1 и 2. Это показывает, что включение уже одной из двух макропеременных почти полностью учитывает влияние изменяющегося макроокружения.

Модель 24 включает переменную CRISIS, которая принимает значение 0 до кризиса августа 1998 г. и 1 – после кризиса. Поскольку графики переменных ERATE и CRISIS схожи, неудиви-

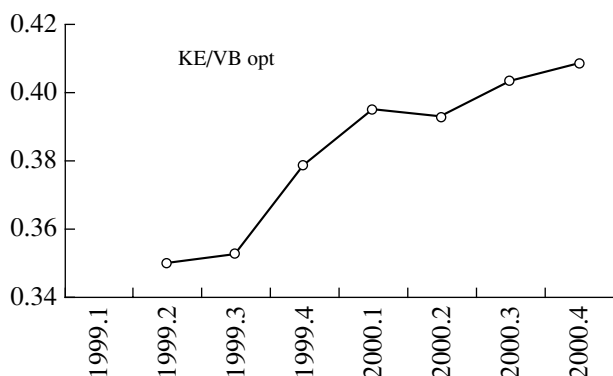


Рис. 9. Расчетное "оптимальное" значение KE/VB.

Эти значения меньше, чем обычно рекомендуемое значение, равное доле 1 в выборке (в нашем случае 0.92).

Знак коэффициента при переменной KE/VB в модели 21 (табл. 17) равен знаку выражения $2.295 + 0.865 \times \text{EXP/IMP}$, что указывает на то, что влияние доли кредитов экономике может быть положительным, если соотношение экспорт-импорт $\text{EXP/IMP} > 2.655$. Однако на протяжении рассматриваемого периода (январь 1996 – апрель 2001) это соотношение не превосходило 2.5.

4. ДИНАМИКА ИЗМЕНЕНИЙ БАНКОВСКОЙ СИСТЕМЫ ПОСЛЕ КРИЗИСА

В этом разделе рассматривается динамика изменений в российской банковской системе. Используется скользящая регрессия с окном шириной в 4, 6 и 8 кварталов. В каждом окне оценивается probit-модель вероятности дефолта. Поскольку набор данных, по которому оценивается модель, теперь гораздо меньше, чем набор данных, используемый в разд. 3, рассматриваются два варианта построения выборки. В первом в выборку включаются все имеющиеся данные. Как и ранее, результаты, полученные по объединенной (pooled) модели, практически не отличаются от результатов модели панельных данных со случайным эффектом. Во втором варианте применяется процедура прореживания данных, отличающаяся от описанной в п. 3.1 тем, что для каждого банка наблюдения берутся с шагом в один год, а не в два года, как ранее.

Различные комбинации размера окна и выборки дают сходные результаты. В табл. 19 приведены результаты оценивания объединенной probit-модели со скользящим окном шириной 8 кварталов. Квартал в строке "Время" соответствует началу скользящего окна.

Из таблицы видно наличие структурного изменения в банковской системе после кризиса 1998 г. (третий квартал 1998 г.). Переменные BP/VB, KE/VB и PZS/VB, которые были не значимыми до кризиса, стали значимыми после кризиса. Это может означать, что после кризиса банки стали уделять больше внимания финансированию реального сектора экономики. Это может быть также указанием на то, что качество данных балансовых отчетов, особенно по просроченным кредитам (PZS) и доходности (BP), улучшилось с 2000 г. В конце 1990-х гг. многие банки использовали различные схемы для "улучшения" балансовых отчетов, например, "оптимизацию" налогов и сокрытие данных по просроченным кредитам (см. (Soest et al., 2003; Пересецкий и др., 2004)).

В периодах (окнах) после второго квартала 1999 г. значимы обе переменные KE/VB и $(\text{KE/VB})^2$. Это позволяет рассчитать по модели "оптимальное" значение доли инвестиций в ре-

тельно, что и критерии качества этих моделей будут близки. Таким образом, остается открытым вопрос о том, имеется ли влияние макропеременных, или результатом банковского кризиса была структурная перестройка российской банковской системы после кризиса. Этот вопрос обсуждается в разд. 4, в котором рассматривается эволюция российской банковской системы после кризиса.

В табл. 18 приведены оптимальные по эвристическим критериям (с процентной ставкой $r = 15\%$) значения порога s^* для трех моделей из табл. 13.

Оптимальные значения порога лежат в диапазоне 0.834–0.884, а среди моделей с макропеременными диапазон еще меньше (0.834–0.853).

Таблица 18. Оптимальные значения порога

Критерий	Базисная модель	Макромодель 1	Макромодель 2
$PR_U (r = 0.15)$	0.884	0.853	0.848
$PR_P (r = 0.15)$	0.869	0.848	0.834

Таблица 19. Probit-регрессия со скользящим окном*

Параметры	1996.2	1996.3	1996.4	1997.1	1997.2	1997.3	1997.4	1998.1	1998.2	1998.3	1998.4	1999.1	1999.2	1999.3	1999.4	2000.1	2000.2	2000.3	2000.4
BP/VB	-0.03	-1.03	-1.79	-0.16	-0.36	-0.32	-0.08	0.91	1.26	1.54	2.11	1.89	2.16	2.02	1.77	5.83	4.84	4.86	5.93
KE/VB	-1.54	-2.09	-1.21	-0.92	-1.26	-1.10	-1.02	-1.49	-0.72	0.01	0.30	1.30	2.04	2.48	3.51	3.52	4.42	4.23	4.39
(KE/VB) ²	0.93	1.73	1.08	0.95	1.34	1.21	1.16	1.85	1.07	-0.06	-0.83	-1.95	-2.91	-3.51	-4.63	-4.44	-5.62	-5.23	-5.36
NDO/VB	-2.02	-2.41	-1.70	-1.50	-1.44	-1.40	-1.30	-1.31	-1.03	-0.71	-1.10	-1.03	-1.03	-1.16	-1.05	-0.76	-0.76	-0.65	-0.60
PNA/VB	-1.80	-1.79	-1.39	-1.08	-1.00	-1.02	-0.83	-0.55	-0.26	-0.22	-0.39	-0.33	-0.40	-0.35	-0.42	0.06	0.08	0.42	0.69
ln(VB/Defl)	0.33	0.41	0.33	0.33	0.32	0.25	0.16	0.07	0.05	-0.09	-0.01	-0.19	-0.20	-0.25	-0.34	-0.32	-0.50	-0.29	-0.75
ln(VB/Defl) ² × 10	-0.17	-0.20	-0.16	-0.16	-0.16	-0.14	-0.10	-0.06	-0.04	0.02	-0.02	0.06	0.07	0.10	0.14	0.13	0.19	0.10	0.27
PZS/VB	2.84	1.64	1.62	0.82	0.67	-0.63	-0.38	-1.29	-0.73	-0.55	-0.59	-1.78	-1.87	-0.77	-0.96	-1.91	-3.64	-4.57	-6.53
RES/VB	-1.21	-0.35	-0.53	0.25	0.34	1.36	0.81	2.04	0.40	0.40	0.87	1.26	1.15	0.30	-0.03	1.50	1.40	1.81	2.32
ERATE	2.86	0.19	0.05	0.03	0.03	0.03	0.02	0.01	0.00	0.00	0.00	-0.02	0.03	0.11	-0.04	-0.01	0.031	0.02	-0.03
Const	-16.2	-0.27	0.763	0.763	0.863	1.293	1.86	2.63	2.78	3.40	2.88	4.33	3.08	0.87	5.41	4.29	4.33	3.39	7.69
Число наблюдений	1590	2447	3327	4184	5063	5937	6807	6929	6908	6880	6821	6771	6706	6641	6582	6527	6488	6485	5666
МакФадден R ²	0.077	0.103	0.073	0.061	0.066	0.07	0.058	0.044	0.025	0.019	0.029	0.033	0.045	0.057	0.064	0.076	0.092	0.087	0.109

*Выделение полужирным означает значимость на 5%-ном уровне; полужирный курсив – на 10%-ном уровне.

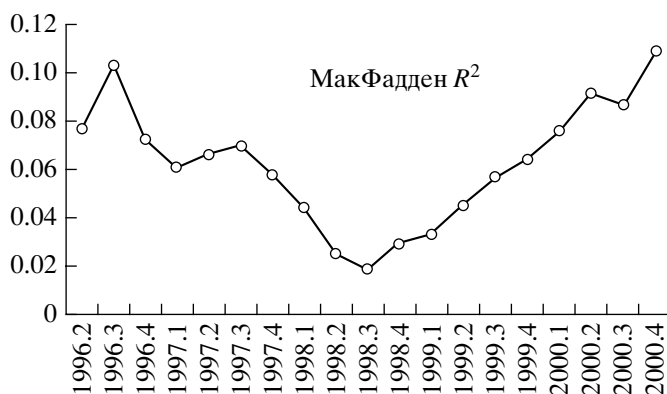


Рис. 10. Изменение качества подгонки модели во времени.

альный сектор экономики в каждом периоде времени. График этого “оптимального” значения представлен на рис. 9. Видно, что за рассматриваемое время это значение возросло с 0.35 до 0.41, что свидетельствует в пользу постепенного улучшения состояния российской банковской системы (поскольку в нормальной экономике финансирование реального сектора является основной функцией банковской системы). Заметим, что отношение суммарных кредитов реальному сектору к суммарным активам банковской системы также возрастало в течение данного периода: 28.6% (1999 г.), 29.5% (2000 г.) и 35.1% (2001 г.).

Обменный курс перестал быть значимой переменной в модели после кризиса, что означает, что в послекризисный период влияние этого макропоказателя на надежность банка серьезно уменьшилось. Волатильность финансовых рынков уменьшилась, и операции стали более прозрачными. Снижился объем высокорискованных финансовых операций.

На рис. 10 приводится эволюция показателя качества подгонки модели МакФадден R^2 во времени. Наименьшее значение соответствует финансовому кризису. Этот показатель постоянно растет после кризиса, указывая как на стабилизацию банковской системы, так и на улучшение качества банковской отчетности.

ЗАКЛЮЧЕНИЕ

В работе получены следующие результаты.

Несмотря на относительно низкое качество данных балансовых отчетов банков, модели вероятности дефолта, основанные на публично доступной информации, могут использоваться в качестве систем раннего предупреждения (EWS).

Предложены модели, учитывающие неоднородность банков. Показано, что их применение улучшает качество прогноза модели.

Показано также, что учет макроиндикаторов улучшает качество прогноза модели.

Результаты процедуры автоматической классификации могут быть использованы органами банковского надзора (ЦБ РФ) для установления более гибких требований к достаточности капитала.

Для использования моделей на практике в меняющемся макроокружении требуется регулярная переоценка параметров моделей.

Предложены эвристические критерии сравнения качества моделей, отражающие точку зрения инвестора.

При помощи оценивания модели со скользящим окном можно прийти к следующим выводам о развитии российской банковской системы после кризиса 1998 г. Тенденция роста качества подгонки модели свидетельствует о стабилизации банковской системы и более предсказуемом макроокружении. Рост значимости в модели таких финансовых показателей банков, как прибыль до вычета налогов и доля просроченных кредитов, говорит об улучшении качества балансовых отчетов. Рост расчетного “оптимального” значения доли кредитов нефинансовым организациям в активах банка указывает на оздоровление банковской системы и рост возможностей для инвестиций в реальный сектор экономики. Данные по суммарным инвестициям банков в реальный сектор экономики показывают, что их рост следует росту расчетного “оптимального” значения с лагом 1–1.5 года.

Модели, подобные предложенным в данной работе, могут быть использованы органами банковского надзора России как элемент системы раннего предупреждения для мониторинга банковской системы, а также коммерческими банками в качестве IRB-подхода к оценке риска в соответствии с соглашением Базель-II.

А. Корреляция финансовых показателей банков, апрель 1998 г.

	LNVB	RES/VB	KE/VB	GDO/VB	SK/VB	LA/VB	VDFL/VB	OS/VB	NDO/VB
LNVB	1	-0.12	0.02	0.18	-0.29	-0.25	0.13	-0.04	0.00
RES/VB	-0.12	1	0.12	-0.16	-0.25	-0.22	-0.03	0.23	-0.07
KE/VB	0.02	0.12	1	-0.22	0.03	-0.25	0.29	-0.15	-0.34
GDO/VB	0.18	-0.16	-0.22	1	0.06	-0.09	-0.04	-0.25	-0.09
SK/VB	-0.29	-0.25	0.03	0.06	1	0.11	-0.09	-0.23	0.24
LA/VB	-0.25	-0.22	-0.25	-0.09	0.11	1	-0.14	-0.28	-0.16
VDFL/VB	0.13	-0.03	0.29	-0.04	-0.09	-0.14	1	0.05	-0.12
OS/VB	-0.04	0.23	-0.15	-0.25	-0.23	-0.28	0.05	1	-0.24
NDO/VB	0.00	-0.07	-0.34	-0.09	0.24	-0.16	-0.12	-0.24	1

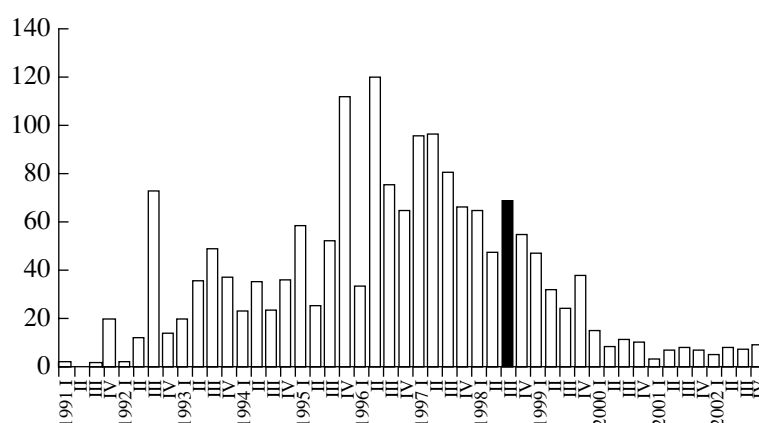
В. Распределение банкротств российских банков, 1991–2002 гг.

Рис. 11. Число банкротств российских банков. Столбец, соответствующий кризису августа 1998 г., выделен черным цветом.

С. Средние значения финансовых показателей банков по кластерам, 1998 г.

Параметры	Число наблюдений	% дефолтов	RES/VB	VDFL/VB	OS/VB	NDO/VB	LA/VB	GDO/VB	KE/VB	SK/VB	LN/VB
Все	1569	17	0.034	0.063	0.202	0.117	0.136	0.073	0.29	0.281	10.72
LIVE = 0	263		0.056	0.049	0.285	0.139	0.073	0.024	0.267	0.174	10.53
LIVE = 1	1306		0.029	0.065	0.185	0.113	0.149	0.083	0.295	0.303	10.76
VB = 1%	624	21	0.039	0.048	0.216	0.11	0.174	0.051	0.282	0.343	8.96
VB (другие)	684	12	0.032	0.073	0.194	0.135	0.123	0.08	0.297	0.265	11.2
VB = 90%	261	18	0.024	0.071	0.189	0.09	0.082	0.11	0.29	0.175	13.67
KE/VB < 15%	392	22	0.021	0.027	0.213	0.199	0.202	0.103	0.062	0.312	10.3
15% < KE/VB < 40%	752	16	0.037	0.068	0.219	0.114	0.121	0.078	0.275	0.25	11.05
KE/VB > 40%	425	13	0.038	0.087	0.161	0.048	0.103	0.037	0.526	0.307	10.52
GDO/VB < 0.01%	624	26	0.043	0.048	0.233	0.127	0.158	0	0.301	0.3	9.6
0.01% < GDO/VB < 10%	567	15	0.03	0.079	0.213	0.12	0.122	0.037	0.315	0.246	11.52
GDO/VB > 10%	378	4	0.024	0.062	0.135	0.099	0.122	0.248	0.234	0.303	11.36
SK/VB < 11%	268	40	0.054	0.045	0.273	0.072	0.112	0.051	0.246	-0.012	11.28
11% < SK/VB < 30%	686	13	0.031	0.076	0.196	0.108	0.134	0.081	0.302	0.205	11.21
SK/VB > 30%	615	10	0.028	0.056	0.178	0.148	0.15	0.074	0.296	0.494	9.92

Д. Оценки модели по разным кластерам, 1998 г.

Модель	m0	m1	m2	m3	m4	m5	m6	m7	m8
Переменная	Все	VB = 1%	VB = 90%	GDO/VB < 0.01%	GDO/VB > 10%	KE/VB < 15%	KE/VB > 40%	SK/VB < 11%	SK/VB > 30%
Const	-7.30***	-0.27	-5.40	-1.80	-9.06	-6.23***	-7.18	-6.35*	-2.47
SK/VB	2.24***	1.65***	6.05***	1.54***	7.11***	1.64***	1.76**	1.85*	-1.34
RES/VB	-5.19***	-5.06***	-8.30	-2.58	-10.84	-7.40*	-1.27	-12.30***	1.03
LNVB	1.42***	-0.19	1.51	0.15	2.42	1.21***	1.73***	1.32***	0.60
LNVB ²	-0.06***	0.03	-0.07	0.005	-0.11*	-0.05**	-0.07***	-0.057***	-0.016
KE/VB	3.27***	4.56***	-4.74	2.80*	-10.59	2.42	-10.22	0.785	2.72
(KE/VB) ²	-4.18***	-5.64***	7.67	-3.19	21.77	-24.09	8.81	0.218	-3.18
GDO/VB	6.07***	8.40***	1.25	12459.2	-0.03	5.86***	10.93**	5.53*	12.7***
LA/VB	3.80***	3.37***	0.13	3.92***	9.38	3.83***	8.82***	3.21**	3.85**
NDO/VB	-1.76***	-1.62*	-5.70***	-0.95	-6.62**	-1.57*	1.23	-1.74	0.16
OS/VB	-1.68***	-1.77**	-2.19	-1.12	-6.55*	-2.12***	0.97	-1.44	-1.80
VDL/VB	1.59	3.94**	-1.99	3.96**	-3.54	1.47	0.98	-0.94	4.22*
МакФадден R ²	0.2	0.22	0.19	0.15	0.28	0.29	0.14	0.26	0.2
Наблюдений с LIVE = 0	263	132	47	161	16	86	57	107	64
Наблюдений с LIVE = 1	1306	492	214	464	362	306	368	161	551
Всего наблюдений	1569	624	261	625	378	392	425	268	615

Автор выражает благодарность А.А. Карминскому и С.В. Голованю за ценные обсуждения и помощь в подготовке данных. Также автор признателен участникам семинаров института BOFIT Банка Финляндии, Тилбургского университета, университета Билкент, РЭШ, ЦЭМИ (семинар С.А. Айвазяна) за полезные комментарии. Автор благодарен А.А. Петрову (агентство “Мобиле”) за предоставленные данные по российским банкам. Эта работа подготовлена во время визита автора в BOFIT, которому автор признателен за предоставленную возможность работы.

СПИСОК ЛИТЕРАТУРЫ

- Бобышев А.А.** (2001): Типичные стратегии и финансовое посредничество. РЭШ. Серия “*Лучшие студенческие работы*”. BSP/01/047.
- Головань С.В., Карминский А.М., Копылов А.В., Пересецкий А.А.** (2003): Модели вероятности дефолта российских банков. I. Предварительное разбиение банков на кластеры. Препринт РЭШ. WP/2003/039.
- Головань С.В., Евдокимов М.А., Карминский А.М., Пересецкий А.А.** (2004): Модели вероятности дефолта российских банков. II. Влияние макроэкономических факторов на устойчивость банков. Препринт РЭШ. WP/2004/043.
- Пересецкий А.А., Карминский А.М., Суст А.Г.О. ван** (2004): Моделирование рейтингов надежности российских банков // *Экономика и мат. методы*. Т. 40. № 4.
- Aldrich J.H., Nelson F.D.** (1985): Linear Probability, Logit and Profit Models. Quantitative Applications in the Social Sciences Series № 45. Beverly Hills: SAGE Publications.
- Altman E.I.** (1968): Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy // *J. of Finance*. Vol. 23. № 4.
- Altman E.I., Marco G., Varetto F.** (1994): Corporate Distress Diagnosis: Comparisons using Linear Discriminant Analysis and Neural Networks (the Italian Experience) // *J. of Banking and Finance*. Vol. 18. № 3.

- Altman E.I., Rijken H.A.** (2004): How Rating Agencies Achieve Rating Stability // *J. of Banking and Finance*. Vol. 28. № 11.
- Amato J.D., Furfine C.H.** (2003): Are Credit Ratings Procyclical? BIS Working Papers. № 129.
- Basel-II** (2004): Basel Committee on Banking Supervision International Convergence of Capital Measurement and Capital Standards. Bank for International Settlements. June 2004. [Http://www.bis.org/publ/bcbs107.htm](http://www.bis.org/publ/bcbs107.htm).
- Borio C.** (2003): Towards a Macprudential Framework for Financial Supervision and Regulation? BIS Working Papers. № 128.
- Borodovsky M., Peresetsky A.** (1994): Deriving Non-homogeneous DNA Markov chain Models by Cluster Analysis Algorithm Minimizing Multiple Alignment Entropy // *Computers and Chemistry*. Vol. 18. № 3.
- Bovenzi J.F., Marino J.A., McFadden F.E.** (1983): Commercial Bank Failure Prediction Models // *Economic Rev.* Vol. 68 (November 1983). Federal Reserve Bank of Atlanta.
- Cole R.A., Gunther J.W.** (1995): Separating the Likelihood and Timing of Bank Failure // *J. of Banking and Finance*. Vol. 19. № 6.
- Cole R.A., Gunther J.W.** (1998): Predicting Bank Failures: A Comparison of On- and Off-site Monitoring Systems // *J. of Financial Services Research*. Vol. 13. № 2.
- Demirgüç-Kunt A., Detragiache E.** (1998): The Determinants of Banking Crises in Developed and Developing Countries // *IMF Staff Papers*. Vol. 45. № 1.
- Engelman B., Porath D.** (2003): Empirical Comparison of Different Methods for Default Probability Estimation. Quanteam Research Paper. [Http://www.quanteam.de/publications.html](http://www.quanteam.de/publications.html).
- Espahbodi H., Espahbodi P.** (2003): Binary Choice Models and Corporate Takeover // *J. of Banking and Finance*. Vol. 27. № 4.
- Estrella A., Park S., Peristiani S.** (2000): Capital Ratios as Predictors of Bank Failure // *FRBNY Econ. Policy Rev.* Vol. 6. № 2.
- Godlewski C.** (2004): Are Bank Ratings Coherent with Bank Default Probabilities in Emerging Market Economies? SSRN. [Http://ssrn.com/abstract=588162](http://ssrn.com/abstract=588162).
- Gunther J.W., Moore R.R.** (2003): Early Warning Models in Real Time // *J. of Banking and Finance*. Vol. 27. № 10.
- Jagtiani J., Kolari J., Lemieux C., Shin H.** (2003): Early Warning Models for Bank Supervision: Simpler Could be Better // *Econ. Perspectives*. Vol. 27. № 3. Federal Reserve Bank of Chicago.
- Kolari J., Glennon D., Shin H., Caputo M.** (2002): Predicting Large US Commercial Bank Failures // *J. of Econ. and Business*. Vol. 54. № 4.
- Komulainen T., Lukkarila J.** (2003): What Drives Financial Crises in Emerging Markets? // *Emerging Markets Rev.* Vol. 4. № 3.
- Korobow L., Stuhr D.P.** (1983): The Relevance of Peer Groups in Early Warning Analysis // *Econ. Rev.* Vol. 68 (November 1983). Federal Reserve Bank of Atlanta.
- Lawrence C.L., Smith L.D., Rhoades M.** (1992): An Analysis of Default Risk in Mobile Home Credit // *J. of Banking and Finance*. Vol. 16. № 2.
- Lennox C.** (1999): Identifying Failing Companies: a Reevaluation of the Logit, Probit and DA Approaches // *J. of Econ. and Business*. Vol. 51. № 4.
- Löffler G.** (2004): An Anatomy of Rating Through the Cycle // *J. of Banking and Finance*. Vol. 28. № 3.
- Martin D.** (1977): Early Warning of Bank Failure: A Logit Regression Approach // *J. of Banking and Finance*. Vol. 1. № 3.
- Mathe C., Peresetsky A., Dehais P., Montagu van M., Rouze P.** (1999): Classification of Arabidopsis Thaliana Gene Sequences: Clustering of Coding Sequences into Two Groups According to Codon Usage Improves Gene Prediction // *J. of Molecular Biology*. Vol. 285. № 5.
- Marchesini R., Perdue G., Bryan V.** (2004): Applying Bankruptcy Prediction Models to Distressed High-yield Bond Issues. // *J. of Fixed Income*. Vol. 13. № 4.
- Ohlson J.A.** (1980): Financial Ratios and the Probabilistic Prediction of Bankruptcy // *J. of Accounting Res.* Vol. 18. № 1.
- Peresetsky A., Karminsky A., Golovan S.** (2004): Probability of Default Models of Russian Banks. Bank of Finland BOFIT Discussion Paper № 21/2004.
- Sahajwala R., Bergh van den P.** (2000): Supervisory Risk Assessment and Early Warning Systems. Basel Committee on Banking Supervision. Working Paper 4.
- Scott A.J., Wild C.J.** (1986): Fitting Logistic Models under Case-control or Choice-based Sampling // *J. of the Royal Stat. Society. Series B*. Vol. 48. № 2.
- Segoviano M.A., Lowe P.** (2002): Internal Ratings, the Business Cycle and Capital Requirements: Some Evidence from an Emerging Market Economy. BIS Working Papers. № 117.
- Soest van A.H.O., Peresetsky A.A., Karminsky A.M.** (2003): An Analysis of Ratings of Russian Banks. Tilburg University Center. Discussion Paper. Series. № 2003/85.

- Stone M., Rasp J.** (1991): Tradeoffs in the Choice between Logit and OLS for Accounting Choice Studies // *Accounting Rev.* Vol. 66. № 1.
- Wescott S.H.** (1984): Accounting Numbers and Socioeconomic Variables as Predictors of Municipal General Obligation Bond Ratings // *J. of Accounting Res.* Vol. 22. № 1.
- Westgaards S., Wijst van der N.** (2001): Default Probabilities in a Corporate Bank Portfolio: A Logistic Model Approach // *European J. of Operational Res.* Vol. 135.
- Wiginton J.C.** (1980): A Note on the Comparison of Logit and Discriminant Models of Consumer Credit Behaviour // *J. of Financial and Quantitative Analysis.* Vol. 15. № 3.

Поступила в редакцию
12.01.2007 г.

Probability Default Models for the Banks

A. A. Peresetsky

The paper presents the results of an econometric analysis of the Russian banks' defaults during the period of 1997–2003, focusing on the extent to which publicly available information from quarterly banks' balance sheets is useful in predicting future banks' defaults. Binary choice models are estimated to construct the probability of default model. We find that preliminary expert clustering or automatic clustering improves the predictive power of the models and incorporation of macrovariables into the models is useful. Heuristic criteria are suggested to help compare model performance from the perspectives of investors or banks supervision authorities. Russian banking system trends after the crisis 1998 are analyzed with rolling regressions.