

Прикладная эконометрика, 2017, т. 48, с. 22–43.
Applied Econometrics, 2017, v. 48, pp. 22–43.

Е. В. Румянцева, К. К. Фурманов¹

Реализация ипотечного имущества: анализ с помощью моделей времени жизни

В статье по данным агентства ипотечного кредитования изучаются сроки реализации недвижимости, заложенной по дефолтным кредитам, и выявляются признаки, отражающие высокий риск длительной реализации. Согласно полученным оценкам, основными детерминантами времени реализации являются тип жилья под залогом, доля заемных средств в стоимости жилья и принадлежность к экономически развитым регионам.

Ключевые слова: ипотечное кредитование; реализация имущества; модели времени жизни.

JEL classification: C41; G21.

1. Введение

В литературе, посвященной эконометрическому моделированию процессов, связанных с ипотечным кредитованием, в центре внимания исследователей находятся, как правило, риски кредитного дефолта и досрочного погашения — иначе говоря, риски «недожития» договора до установленного срока. Существует и еще один тип риска ипотечного кредита, который участвует в задаче моделирования денежных потоков по ипотечному портфелю и непосредственно связан с риском наступления дефолта, являясь по отношению к нему вторичным. После дефолта начинается процесс реализации имущества, находящегося под залогом дефолтного кредита, в рынок с торгов вследствие судебных процедур или после промежуточной стадии постановки залога на баланс кредитора с последующей реализацией третьим лицам. Этот процесс растянут во времени и связан с возможностью денежных потерь в случаях, когда реализовать имущество в короткие сроки не удастся.

Оценивание риска реализации составляет основу определения величины потерь при дефолте ипотечного кредита. Динамика реализации объектов с момента наступления дефолта является определяющей в длительности процессов взыскания кредитора с ипотечного должника, поэтому важно иметь оценки временной зависимости² риска реализации залоговой недвижимости. Знание временной зависимости позволяет также определить объем совокупного требования по дефолтному кредиту к наиболее вероятному моменту реализации

¹ Румянцева Екатерина Владимировна — ОАО «Агентство по ипотечному жилищному кредитованию», Москва; Evrumyantseva.2006@yandex.ru.

Фурманов Кирилл Константинович — Национальный исследовательский университет «Высшая школа экономики», Москва; furmach@rbcmail.ru.

² Временная зависимость (duration dependence) — зависимость вероятности наступления события от времени его ожидания.

залогового имущества и, соответственно, определить остаточный объем ожидаемых потерь, не покрываемых залогом.

На текущий момент нет опубликованных исследований времени реализации ипотечного залога в России, и цель настоящей статьи — восполнить этот пробел. В статье приводятся основные статистические показатели процесса реализации, а также результаты эконометрического моделирования этого процесса.

Методологическую основу исследования составляют регрессионные модели времени жизни (survival models), позволяющие полностью описать распределение времени до реализации и, как следствие, оценить вероятность реализации в течение какого-либо срока.

Интерес исследователей к эконометрическому моделированию процессов, связанных с ипотечным кредитованием, во многом обусловлен влиянием последнего ипотечного кризиса в США 2006–2008 гг. Тогда сроки экспозиции для продажи залоговой недвижимости достигали исторических максимумов, и ее характеристики (как и готовность клиента способствовать процессу) играли определяющую роль в скорейшем завершении процесса реализации.

Низкий интерес к моделированию реализации ипотечного залога в России может объясняться, с одной стороны, относительно невысокой долей ипотечных сделок в общем объеме сделок купли-продажи недвижимости по сравнению с развитыми странами (например, США и Великобританией), а с другой стороны, низкоконкурентным в целом ипотечным рынком в России, где на долю пяти крупнейших банков в рознице приходится практически весь объем сделок. Тем не менее, развитие современных подходов к оценке денежных потоков по ипотечному кредиту во всех его аспектах (риск дефолта, риск досрочного погашения и вероятности реализации залога) способно привлечь интерес инвесторов к ипотечным активам — в том числе посредством спроса на секьюритизированные продукты — и придать импульс развитию ипотечного рынка. Ввиду внедрения в 2017 г. стандартов МСФО 9 оценка ипотечных кредитов с учетом всего времени жизни кредита приобретает и регуляторный характер, что означает потребность в надежных временных оценках риска.

В данной статье предлагаются оценки вероятности реализации для грамотного построения денежных потоков по пулу ипотечных кредитов в части описания доли денежных средств, возвращенных кредитору вследствие продажи имущества. Расчет времени до реализации залогового имущества позволяет в каждый момент времени произвести оценку величины ожидаемых потерь кредитора при дефолте ипотечного кредита. Кроме того, задача оценки времени до реализации залога сама по себе является важным направлением исследования в понимании причин и факторов, влияющих на скорость реализации имущества.

2. Обзор литературы

Работы, связанные с темой настоящего исследования, в основном посвящены оцениванию функций, участвующих в задаче моделирования денежных потоков по портфелю кредитов, описывающих ставку восстановления заемных средств кредитору (recovery rate, RR) при дефолте кредита, а также величины потерь при дефолте кредита (loss given default, LGD = 1 – RR). Обе эти функции участвуют в задаче моделирования денежных потоков по портфелю кредитов.

Задача оценки ставки восстановления заемных средств кредитору и, следовательно, суммы денежных средств, возвращенных кредитору (recovery amount) вследствие процедуры обращения взыскания по дефолтному кредиту, стала очень важной из-за увеличения объема дефолтных кредитов в сфере потребительского кредитования в кризис 2007–2008 гг. В соответствии с Базельским соглашением было разрешено рассчитывать капитал под рискованные активы на основании так называемого подхода IRB (internal rating-based approach), основанного на развитии собственных методик по оценке следующих риск-параметров кредитного риска портфеля закладных:

- вероятности дефолта в течение следующих 12 месяцев PD (probability of default);
- потерь при дефолте LGD;
- ожидаемого объема требования к моменту дефолта EAD (expected exposure at default).

В отличие от моделирования вероятности дефолта, являющегося целью различных кредитных скоринговых систем в течение последних пятидесяти лет, техника моделирования LGD начала развиваться только с начала 2000-х гг. Необходимость оценки LGD явилась следствием Базельских рекомендаций. В одной из первых статей (Schuermann, 2004), посвященных данной теме, приводится обзор того, что было известно относительно LGD к моменту исследования. С момента появления Базельских консультативных работ отмечается возросшее число исследований, посвященных разным техникам оценки LGD, таких как (Matuszyk et al., 2010; Altman et al., 2004; Gupta, 2005; Huang, Oosterlee, 2008).

Например, в работе (Matuszyk et al., 2010) для моделирования величины потерь при дефолте LGD по необеспеченным потребительским кредитам применяется подход, основанный на дереве решений (tree decision approach). По мнению авторов, процесс обращения взыскания долга в пользу кредитора может проводиться тремя способами:

- 1) собственными силами банка;
- 2) с использованием агента и выплатой ему соответствующей комиссии, как правило, равной доли от суммы средств, полученных в результате возврата долга кредитору;
- 3) путем продажи долга третьим лицам по фиксированной цене.

Данные стратегии подразумевают различные диапазоны значений величины потерь LGD. Так, если при проведении процедуры возврата долга самим банком величина LGD может принимать любое значение в диапазоне от 0 до 1, то в случае привлечения агента эта величина может варьироваться в пределах от 0.4 до 1 при условии 40% комиссии. При продаже долга третьим лицам, например по цене, равной 5% от номинальной стоимости, величина потерь LGD составляет 0.95. В результате исход представляет собой смесь решений и случайностей, для моделирования которых применяется дерево решений.

Построив LGD распределение на опытных данных по дефолтным банковским кредитам при условии, что процедура возврата долга производится самим банком и остаточный невыплаченный долг в конце концов списывается кредитором, авторы увидели, что 30% должников полностью погасили долг ($LGD = 0$), а менее 10% заемщиков не заплатили ничего. Для некоторых должников LGD превышает 1, это может означать, что к первоначальной сумме требования по дефолтному кредиту были добавлены пени, штрафы и т. д. Подобное распределение предполагает существование двух групп должников — тех, которые способны оплатить долг полностью, и тех, кто не может выплатить требуемую сумму долга. Это соответствует разбиению данных на множества: $\{LGD = 0\}$ и $\{LGD > 0\}$. В других случаях статистический анализ и отчасти понимание политики взыскания кредитора позволяют провести более детальную разбивку должников на относительно однородные группы по уровню потерь.

Поэтому при моделировании LGD авторы применяют двухшаговую процедуру, используя характеристики должника и заемщика. На первом этапе используется логистическая регрессия для того, чтобы выделить две группы: $\{LGD = 0\}$ и $\{LGD > 0\}$, или кумулятивная логистическая регрессия, если требуется выделить более двух подмножеств. На втором шаге внутри каждой группы строятся модели регрессионного типа для того, чтобы оценить LGD для каждого кредита. Понятно, что для каждого наблюдения из первого класса предсказанное значение LGD будет равно 0. Для второй группы LGD тестируются различные подходы, но лучшей оказывается модель линейной регрессии, основанная на всей совокупности данных.

Моделирование потерь при дефолте является более сложной задачей, чем моделирование вероятности дефолта, по двум причинам. Во-первых, большинство данных цензурируется из-за длительности процесса восстановления. Поэтому линейная регрессия не годится для цензурированных данных. Во-вторых, должники имеют различные причины, приводящие к дефолту, и учет данного фактора приводит к различным моделям погашения. Например, одни заемщики не хотят погашать долг. Некоторые должники не могут его погасить вследствие постоянного изменения жизненных обстоятельств, в то время как для других причины непогашения временные. Все это многообразие ситуаций невозможно описать единым распределением. Однако отмеченные особенности в данных позволяют учесть модели времени жизни, которые как раз и подходят для цензурированных данных, а удачное сегментирование дефолтного множества кредитов позволяет описать разнообразие причин, приведших к дефолту, при моделировании LGD.

В части моделирования ставки восстановления долга RR и величины потерь при дефолте LGD методами анализа дожития выделим для относительно подробного рассмотрения исследования (Zhang, Thomas, 2012) и (Witzany et al., 2012). В них оценивание проводится на основании данных по необеспеченным (незалоговым) кредитам.

В работе (Zhang, Thomas, 2012) для прогнозирования величины RR и, следовательно, LGD авторы используют модели линейной регрессии и модели времени жизни. Целью проводимого анализа является сравнение моделей между собой и выбор наилучшей. Исследование проводится на основании данных по потребительским кредитам, которые являются дефолтными и находятся в процессе обращения взыскания — возврата заемных средств кредитору. Авторы оценивают линейные регрессии стандартного вида, где объясняемыми переменными выступают темп восстановления RR и сумма долга, которую удалось взыскать. Величина RR рассчитывается как отношение суммы долга, которую удалось взыскать, к сумме требования на момент дефолта кредита. В качестве объясняющих переменных используются характеристики кредита или должника (заемщика). Авторы отмечают, что распределение случайной ошибки существенно отличается от нормального, что, по их мнению, служит доводом в пользу моделей времени жизни.

В действительности, термин «модели времени жизни» употребляется в работе условно, т. к. объясняемая величина отражает не «время жизни» до полного взыскания, а все ту же ставку восстановления или сумму взысканного долга. Авторы употребляют понятийный аппарат и инструментарий анализа дожития, потому что сталкиваются с проблемами, типичными для этой дисциплины: не нормальным распределением случайной ошибки и цензурированием. Последнее возникает, поскольку к моменту среза данных не для всех кредитов завершается процедура возврата долга. Авторы указывают на возможность использования цензурированных данных как на существенное преимущество методов анализа дожития,

поскольку можно использовать данные по возврату долга, не дожидаясь, пока он будет полностью погашен заемщиком или списан банком-кредитором. При сравнении моделей времени жизни с линейной регрессией оказывается, что последняя лучше при сравнении среднеквадратической ошибки прогноза и степени корреляции прогнозов и наблюдаемых значений, измеренной с помощью R^2 и рангового коэффициента Спирмена. Авторы, однако, делают вывод, что для объяснения ставки восстановления долга RR нельзя отказаться от моделей времени жизни в пользу простой линейной регрессии по причине серьезного истощения выборки, возникающего при отказе от цензурированных наблюдений.

Работа (Witzany et al., 2012) также посвящена использованию моделей времени жизни для оценки LGD, но при этом слова «время жизни» имеют обычный смысл: объясняемой величиной в оцениваемых моделях выступает время до полного погашения долга. Методология имеет интересные особенности: вместо традиционных методов (максимального правдоподобия и частичного правдоподобия Кокса) оценивание опирается на минимизацию суммы квадратов ошибок с поправкой на цензурирование. Исследование проводилось на основании LGD-данных по 4000 необеспеченным дефолтным закладным крупного чешского ритейл-банка.

В работе авторы вводят понятия реализовавшегося и ожидаемого темпа восстановления (RR) и связанного с ними понятия потерь при дефолте LGD. Реализовавшийся RR можно наблюдать только по случаям, вошедшим в дефолт, в то время как ожидаемый RR оценивается по недефолтным случаям на основании доступной информации. Показатели RR и LGD ($LGD = 1 - RR$) рассчитываются как доли от величины EAD — суммы требования к моменту дефолта кредита — и обычно считаются меньшими 1. Далее для расчета RR учитывается чистый денежный поток по наблюдениям, находящимся в процессе обращения взыскания (возврата долга). Полученный денежный поток суммируется и дисконтируется по ставке, отражающей риск в соответствии с регуляторными требованиями Базельского комитета по риску (BCBS, 2006). Тогда показатель RR рассчитывается как отношение полученной в результате величины к сумме требования на момент дефолта. После оценки реализовавшегося темпа восстановления RR следующей задачей является расчет LGD по недефолтным закладным. Для выдачи нового кредита аппликационный скоринг, проводимый банком, должен учесть не только вероятность дефолта в течение 12 или более месяцев после выдачи кредита, но и LGD на том же временном горизонте. При расчете кредитной ставки кредитная маржа должна покрывать, помимо стоимости фондирования, административных издержек и др., еще и ожидаемые потери в случае дефолта кредита. Ожидаемый LGD также должен рассчитываться банками, применяющими подход IRB (internal rating-based approach), чтобы вычислить требование капитала для каждого недефолтного наблюдения в соответствии с Базельскими рекомендациями. Ожидаемый LGD авторы рассчитывают с помощью множества объясняющих переменных, учитывая специфические свойства каждой недефолтной закладной, основываясь на исторических данных по темпам возврата долга RR и наблюдаемым значениям объясняющих переменных. В результате исследования оказалось, что предложенная авторами полупараметрическая модель (pseudo Cox model) лучше описывает данные по восстановлению долга, чем параметрические регрессии. Авторы указывают на низкую объясняющую силу линейных регрессий при описании потерь при дефолте и отмечают, что любое возможное повышение точности прогноза оказывается важным для банка — именно модели дожития и предлагаются как инструмент повышения точности.

В отличие от потребительского кредитования, для ипотечного кредитования характерно высокое влияние стоимости залога недвижимости, по которой происходит обращение взыскания по дефолтному кредиту, на величину потерь при дефолте. Так, необходимо оценить время до реализации залога третьим лицам в счет погашения задолженности, от которого зависит величина необеспеченного требования по кредиту: чем дольше происходит продажа недвижимости, тем больше накапливается сумма требования кредитора по кредиту и тем больше становится остаточная величина необеспеченного остатка задолженности после учета продажной стоимости залоговой недвижимости. Зачастую остаточное требование после реализации залога практически не погашается заемщиком и составляет основную часть величины потерь при дефолте (LGD). Таким образом, для ипотечного кредитования оценка вероятности реализации залога в зависимости от срока после дефолта играет ключевую роль в оценке потерь при дефолте.

Работы по моделированию LGD в розничном кредитовании с применением моделей времени жизни посвящены беззалоговому кредитованию (потребительское кредитование, с использованием кредитных карт и др.). Поэтому настоящая работа, посвященная моделированию вероятности реализации залога на основе данных по ипотечному портфелю кредитов, является новой.

Существует ряд эконометрических исследований ипотечного кредитования в России, из недавних работ можно указать (Лозинская, Ожегов, 2014; Карминский и др., 2015; Lozinskaia, 2015; Румянцева, Фурманов, 2016). В них, в основном, рассматриваются риски дефолта, а в последней статье — дефолта и предварительного погашения. Работа (Карминский и др., 2015) посвящена оценке ожидаемых потерь, однако временному аспекту внимание не уделяется.

В методическом аспекте интересна статья (Козлов, Гушин, 2017), где рассматривается применение моделей времени жизни для анализа рисков в розничном кредитовании, при этом наряду с классическим подходом (оценка Каплана–Майера, регрессия Кокса) используются методы машинного обучения. Авторы настоящего исследования, однако, решили не применять машинное обучение, ограничившись классическим подходом, т. к. он позволяет получать интерпретируемые результаты.

Исследования, содержащие анализ времени до реализации заложенного имущества по российским данным хотя бы на описательном уровне, авторам не известны.

3. Данные

Исследование проводится на основании региональных данных по портфелю крупного российского ипотечного агентства. Исследуется срез данных на 01.01.2016, включающий наблюдения с начала деятельности предприятия в 2001 г. Кредиты выдавались практически во всех регионах России первичными кредиторами — региональными банками или специализированными ипотечными агентствами при региональных правительствах субъектов РФ.

Согласно определению агентства, дефолт по кредиту наступает, если заемщик не производит обязательные ежемесячные платежи в течение не менее трех месяцев. После присвоения кредиту дефолтного статуса с некоторым временным лагом (зависящим от внутренних процедур банка) запускается процедура обращения взыскания имущества, находящегося

под залогом, через суд, т. е. процесс судебной реализации объекта недвижимости, находящегося под залогом дефолтного кредита, с торгов.

При невозможности реализации объекта в рынок с торгов объект недвижимости, как правило, ставится на баланс банка-кредитора. В дальнейшем специальные подразделения занимаются его продажей с баланса. Этот процесс может быть длительным, но все же он предположительно конечен, в отличие от процессов ожидания досрочного погашения и дефолта, часто моделируемых в исследованиях ипотеки. Причина случаев, когда время реализации не отражено в данных, заключается скорее в конечности времени наблюдения за кредитом, в ходе которого не удается «дождаться» интересующего момента, чем в наличии принципиально нереализуемых объектов недвижимости.

Реализация залога наступает, если объект недвижимости продается на рынке с торгов или ставится на баланс банка-кредитора вследствие невозможности его продажи в разумные сроки и впоследствии реализуется. Стоит подчеркнуть, что исследование времени до постановки на баланс, а не времени до реализации залога с баланса банка, производится исключительно вследствие невозможности получения необходимых данных, поскольку подобная информация не является публичной. Таким образом, момент реализации — это либо момент постановки на баланс банка, либо момент продажи с торгов, причем последнее случается редко.

Далее дефолтные кредиты, находящиеся под судом, на основании которых производится исследование, будем ради краткости называть просто дефолтными кредитами.

В настоящей работе исследуется время жизни дефолтного кредита с первого дня просрочки до момента реализации объекта недвижимости, находящегося под залогом, на рынке в зависимости от факторов, свойственных определенной группе кредитов. Объектом исследования являются дефолтные кредиты разной глубины просрочки (но не менее трех периодов), по которым уже проводятся судебные процедуры. Вообще говоря, возможна ситуация, когда кредит, которому присвоен дефолтный статус, возвращается в платежеспособное состояние, т. е. состояние без просрочки или просрочки менее трех платежных периодов. Однако подобные случаи, во-первых, единичны, а во-вторых, рассматриваются банком-кредитором в частном порядке, и решение не всегда бывает в пользу заемщика. Поэтому данное исследование базируется на множестве «безнадежно дефолтных кредитов», одновременно являющихся и дефолтными, и находящимися «под судом». Анализируемые данные содержат 10681 наблюдение за такими кредитами.

Помимо сроков реализации, наблюдения отражают характеристики основного заемщика (на которого оформлен договор), кредитного договора и предмета ипотеки. Для описания портрета основного заемщика отобраны следующие признаки: семейный статус, уровень образования, вид занятости, доля дохода основного заемщика в совокупном доходе созаемщиков. Параметры кредита включают кредитную ставку, срок кредита, а также распространенный в различных ипотечных исследованиях показатель LTV, характеризующий долю заемных средств в стоимости жилья. Для описания жилого помещения, находящегося под залогом (предмета ипотеки), используются следующие переменные: регион предмета ипотеки и тип жилья. Для анализа также сконструирован показатель, рассчитываемый как отношение совокупных денежных выплат за весь срок жизни кредита (плановый срок по графику) к стоимости предмета ипотеки (переменная *payed_sum* в Приложениях). Показатель вычисляется по формуле: $(\text{Аннуитет} \times \text{Срок}) / \text{Стоимость}$. Этот показатель в простейшем смысле показывает переплату за жилье, т. е. характеризует психологическую готовность заемщика выполнять свои кредитные обязательства.

Важной детерминантой времени реализации может выступать заявленная продавцом цена реализуемого имущества, но в настоящем исследовании она не рассматривается, т. к. в большинстве рассматриваемых случаев реализация осуществлялась банками, так что цена формировалась уже после наступления моделируемого события (постановки на баланс банка). Кроме того, в случаях продаж непосредственно с торгов цена на момент вхождения в дефолт также отсутствует, и поэтому не может определять вероятность продажи в течение значительной доли исследуемого временного промежутка.

Все характеристики были представлены бинарными переменными, использовались их значения на момент выдачи кредита. Выбор пороговых значений при категоризации количественных характеристик связан с определенными кредитными программами агентства. Распределение переменных приведено в Приложении 1.

4. Методология

Методологическую основу исследования составляют статистические методы анализа дожития, учитывающие наличие цензурирования в данных: для некоторых кредитов реализация не наступает за время наблюдения, поэтому точное время жизни таких кредитов неизвестно. Под временем жизни здесь понимается длительность периода от попадания кредита в категорию дефолтных до реализации.

В рамках анализа дожития распределение исследуемой величины (времени до наступления целевого события) задается, как правило, с помощью функций дожития, риска и интегрального риска, определяемых следующим образом (Kiefer, 1988). Пусть время до реализации описывается неотрицательной случайной величиной T с функцией плотности $f(t)$. *Функция дожития* величины T задается соотношением

$$S(t) = P(T > t) = \int_t^{\infty} f(x) dx.$$

Значение функции дожития равно вероятности того, что реализация имущества потребует более t единиц времени.

Функция риска (hazard function) отражает «склонность к реализации в ближайшие сроки» для имущества, ожидающего реализацию уже t единиц времени:

$$h(t) = \lim_{dt \rightarrow 0} \frac{P(t < T \leq t + dt | T > t)}{dt} = \frac{f(t)}{S(t)}.$$

Если функция риска возрастает, то говорят, что существует положительная временная зависимость (duration dependence) — с течением времени вероятность реализации в ближайшие сроки увеличивается. При убывании функции риска говорят об отрицательной временной зависимости.

Отдельный интерес представляет интегральная (кумулятивная) функция риска

$$H(t) = \int_0^t h(x) dx.$$

По ее виду также можно судить о характере временной зависимости: интегральная функция риска выпукла вверх на участках положительной временной зависимости и выпукла вниз при отрицательной. У величин, распределенных по показательному закону, интегральная функция риска линейна — это используется при визуальной проверке спецификации регрессионных моделей дожития.

В настоящей статье для изучения сроков реализации последовательно применяются следующие подходы, составляющие методологическую основу четырех шагов эмпирической части исследования.

- *Одномерный анализ* — описание функций дожития и риска вне связи с какими-либо переменными.

- *Сравнение описательных статистик* для характеристик кредита, заемщика и заложенного жилья во всей выборке и в выборке «проблемных» кредитов. «Проблемными» условно назывались те кредиты, для которых время реализации оказалось больше медианы. Поскольку точная дата реализации известна не всегда, то в выборку «проблемных» включались лишь те наблюдения, у которых время реализации либо известно и оказалось больше медианного, либо не известно, но на момент сбора данных кредит ожидал реализации уже не менее медианного времени.

- *Модели с объясняющими переменными* для изучения связи срока реализации с характеристиками кредита, заемщика и заложенного имущества.

- *Модели разделенной совокупности* (split-population models), позволяющие выделить латентный класс «безнадежных» объектов — тех, для которых время реализации можно условно считать бесконечным.

Одномерный анализ опирается на непараметрические методы оценивания распределения времени жизни: оценку Каплана–Майера для функции дожития (Kaplan, Meier, 1958) и Нельсона–Аалена для интегральной функции риска (Nelson, 1972; Aalen, 1978) — эти методы позволяют дать общую характеристику процесса реализации. Кроме того, на основании оценки Каплана–Майера рассчитывается выборочная медиана времени реализации, используемая при сравнении описательных статистик на втором шаге исследования.

Для изучения связи срока реализации с объясняющими переменными в статье используются стандартные регрессионные модели пропорциональных рисков и ускоренного времени, а также их модификации с учетом ненаблюдаемой разнородности.

Модели пропорциональных рисков основаны на предположении, что вектор регрессоров x мультипликативно входит в функцию риска (Cox, 1972):

$$h(t; x, \beta) = h_0(t) \exp(x' \beta), \quad (1)$$

где $h_0(t)$ — опорная функция риска (baseline hazard), отражающая распределение длительностей в случае равенства всех регрессоров нулю, β — вектор коэффициентов при объясняющих переменных.

В зависимости от того, какие предположения делаются относительно базового риска $h_0(t)$, модели пропорциональных рисков можно разделить на два вида.

Параметрические регрессии, предполагающие определенную функциональную форму опорного риска. Как правило, используются функции, соответствующие распределениям Вейбулла, Гомперца или показательному. Параметры этих распределений оцениваются методом максимального правдоподобия.

Полупараметрическая регрессия Кокса, не накладывающая ограничений на функцию риска и оцениваемая методом частичного правдоподобия (partial likelihood). При этом зависимость от объясняющих переменных все же ограничивается формулой (1), что позволяет интерпретировать получаемые оценки.

Потенцированные коэффициенты при регрессорах в моделях пропорциональных рисков показывают, во сколько раз изменяется риск реализации при единичном изменении объясняющей переменной и прочих равных условиях.

В *моделях ускоренного времени* предполагается, что регрессоры мультипликативно связаны с масштабом времени (Cleves et al., 2004):

$$S(t; x) = S_0(t \exp(-x'\beta)), \quad (2)$$

где $S_0(t)$ — опорная функция дожития. Выбор функциональной формы опорного распределения приводит к различным моделям ускоренного времени. В настоящем исследовании использовались логнормальная, логлогистическая и обобщенная гамма-регрессия, оценивание проводилось методом максимального правдоподобия.

Потенцированные коэффициенты моделей ускоренного времени показывают, во сколько раз вырастают все квантили объясняющей переменной и ее математическое ожидание (если оно существует) при увеличении регрессора на единицу при прочих равных условиях. Регрессии, опирающиеся на показательное распределение и распределение Вейбулла, допускают представление как в виде (1), так и в виде (2), выбор параметризации может определяться удобством интерпретации.

Учет ненаблюдаемой разнородности в регрессиях дожития производится с помощью вводимого в модель случайного эффекта v , мультипликативно связанного с функцией риска (даже если объясняющие переменные включаются в духе моделей ускоренного времени):

$$h(t; x, v) = h(t; x)v.$$

Приводимые в настоящей статье результаты получены в стандартном предположении, что величина v имеет гамма-распределение с единичным математическим ожиданием и неизвестной дисперсией — такое предположение позволяет получить удобный вид функции правдоподобия. Выражение «ненаблюдаемая разнородность» (в англоязычной литературе используются термины «frailty» и «unobserved heterogeneity») следует понимать условно: в действительности в модель просто вводится дополнительный параметр, позволяющий улучшить подгонку к данным.

Модели разделенной совокупности вошли в исследовательский обиход благодаря работе (Schmidt, Witte, 1987), посвященной анализу рецидивизма. Их особенность — разделение совокупности моделируемых объектов на два латентных класса, в одном из которых событие (в оригинальной статье — возвращение в тюрьму) наступает по истечении случайного времени, а в другом — не наступает никогда. Иногда первый класс называют классом «кочевых» («movers», т. к. они кочуют из состояния в состояние), а второй класс — классом «оседлых» (stayers). В общем виде модель разделенной совокупности можно представить следующим образом (Ратникова, Фурманов, 2014):

$$S(t) = \delta S_m(t) + 1 - \delta.$$

Здесь $S_m(t)$ — функция дожития для класса «кочевых», а δ — вероятность принадлежности к этому классу, так что $1 - \delta$ — вероятность того, что реализация залога никогда не наступит. Такие случаи могут казаться неправдоподобными, но мы допускаем существование класса «почти безнадежных» кредитов, по которым реализация затягивается на сроки, заметно превышающие сроки наблюдения за анализируемыми кредитами.

Связь с объясняющими переменными в модели разделенной совокупности задается двумя уравнениями. Первое уравнение описывает дожитие в классе «кочевых» и может быть задано любой параметрической спецификацией. В настоящей статье используются те же параметрические модели, что и на предыдущем шаге. Второе уравнение — классифицирующее. Оно задает вероятность принадлежности к тому или иному классу и может быть представлено любой моделью бинарного выбора. В данной статье приводятся результаты, полученные с использованием логит-модели

$$\frac{\delta(z; \gamma)}{1 - \delta(z; \gamma)} = \exp(z' \gamma).$$

Здесь z — вектор объясняющих переменных для классифицирующего уравнения, γ — вектор оцениваемых коэффициентов.

Для проверки правильности спецификации моделей дожития используются остатки Кокса–Снелла. Эти остатки представляют собой значения оцененной согласно диагностируемой модели интегральной функции риска в каждом наблюдении и при правильной спецификации имеют распределение, близкое к показательному с параметром $\lambda = 1$, так что график интегральной функции риска для остатков Кокса–Снелла похож на прямую линию — биссектрису угла, образуемого осями координат.

Подробное описание подходов к моделированию дожития по цензурированным данным можно найти в книге (Klein, Moeschberger, 2005), обзор используемых регрессионных моделей содержится в книге (Cleves et al., 2004), особенности оценивания моделей разделенной совокупности — в (Ратникова, Фурманов, 2014).

5. Результаты

Описанные в предыдущем разделе процедуры были применены к располагаемым данным дважды: один раз — ко всем наблюдениям, другой раз — только к наблюдениям за кредитами, которые стали безнадежно дефолтными в 2009 г. и позже. Это было сделано, чтобы избежать искажения результатов из-за возможных структурных сдвигов, связанных с финансовым кризисом, однако оказалось, что результаты, полученные по всем наблюдениям и по их послекризисной части, практически совпадают. В действительности, число наблюдений за кредитами, вошедшими в дефолт до кризиса 2008 г., невелико. Если взять за начало кризиса 1 мая 2008 г., таких наблюдений окажется 973, что составляет менее десятой части выборки, причем во всех этих случаях реализации наступали не ранее 2009 г. Этот факт не удивителен: портфель ранних договоров кредитования составлялся в период, когда только разрабатывались процедуры обращения взыскания на имущество, находящееся под залогом — судебные реализации стали происходить позже.

Ниже приводятся результаты, полученные при анализе полного набора данных.

5.1. Одномерный анализ

Из 10681 дефолтных кредитов, по которым проводились судебные процедуры, в 4430 случаях наступило целевое событие, т. е. объект недвижимости, находящийся под залогом такого кредита, был реализован или поставлен на баланс. Анализ скорости реализации объектов в выборке позволяет заключить, что в среднем в течение дня реализуются 2.9 объекта недвижимости из 10000.

Таблица 1. Распределение времени реализации

Суммарное время наблюдения, дни	Скорость реализации (реализаций в день)	Число наблюдений	Время дожития по квартилям выборки		
			25%	50%	75%
15265004	2.9×10^{-4}	10681	1188	1909	—

Согласно статистике ипотечной организации, на основании данных которой производится исследование, процесс реализации имущества длительный, поскольку в портфеле представлена, в основном, малорентабельная региональная недвижимость. Описательная статистика (табл. 1) показывает, что для 25% дефолтных кредитов реализация объекта недвижимости из-под залога происходит в течение 1188 дней с момента первого вхождения в просрочку 1 день, т. е. в течение примерно трех лет. Для половины дефолтных кредитов реализация объекта недвижимости происходит в течение пяти лет с момента вхождения в просрочку. Для 75% квантили оценка числа реализаций недоступна из-за цензурирования в данных.

Отметим, что время от вхождения в дефолтное состояние до реализации или постановки на баланс банка несопоставимо с временем от постановки на баланс банка до продажи заложенного имущества, поэтому приводимые в настоящей статье характеристики распределения сроков реализации значительно отличаются от тех значений, что приводятся, например, в статье (Лейбов, 2016), где утверждается, что среднее время продажи жилья с баланса банка составляет 3.7 года.

Для наглядного представления динамики реализаций на рис. 1 отображены оценки функции дожития и интегральной функции риска. По горизонтальной оси откладывается время

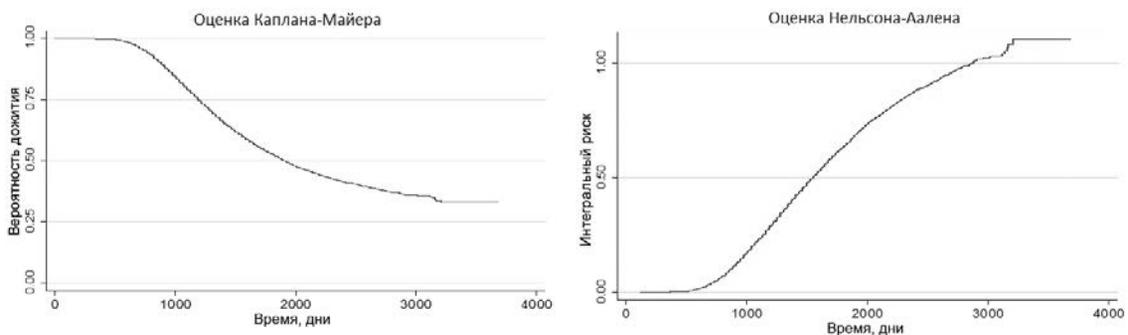


Рис. 1. Оценки для функции дожития (Каплана–Майера) и интегральной функции риска (Нельсона–Аалена)

исследования, а по вертикальной — оцененное значение функции дожития (левый график) или интегрального риска (правый график). Из графика видно, что реализации практически не происходят ранее 600 дней с момента вхождения в просрочку в 1 день. Затем риск реализации начинает расти (интегральная функция выпукла вниз). После примерно 1000 дней риск выходит на стабильный уровень: график оценки Нельсона–Аалена похож на прямую линию, что говорит об отсутствии временной зависимости. Позже риск начинает падать, так что после 3000 дней реализаций почти не наблюдается. Предположительно это связано с существенной разнородностью объектов под залогом: к этому времени остается только проблемное имущество, которое трудно реализовать. Согласно оценке Каплана–Майера, такое имущество составляет более четверти всех объектов.

5.2. Сравнение описательных статистик в полной выборке и в выборке «проблемных» кредитов

К «проблемным» были отнесены кредиты, для которых время реализации оказалось больше медианы (1909 дней по оценке Каплана–Майера). Такие кредиты составили примерно 26% от общего числа — в отсутствие цензурирования эта доля должна быть равна 50%, но на момент последнего наблюдения многие нереализованные объекты еще не дожили до медианного времени.

Распределение наблюдений по характеристикам кредита, заемщика и жилья для полной и проблемной выборки приведено в Приложении 1. Основное различие между двумя выборками касается занятости заемщиков. Доли госслужащих, военных и наемных работников среди проблемных кредитов примерно в три раза ниже, чем в полной выборке. Основная масса трудно реализуемого имущества принадлежит к так называемой базовой категории, куда входили, по большей части, заемщики, о роде занятости которых не было сведений в базе (например, потому что они не предоставили сведений). В эту категорию входили также предприниматели, пенсионеры и учащиеся, но их доля значительно ниже. В полной выборке на долю базовой категории приходится 66% наблюдений, в то время как среди проблемных кредитов эта доля составляет 88%.

Признаком «опасного» заемщика выступает также наличие высшего образования — в выборке проблемных кредитов доля таких заемщиков на 8 процентных пунктов выше, чем в общей массе наблюдений. Наличие высшего образования по определенной категории дефолтных заемщиков обуславливает их более активную позицию по затягиванию судебного процесса. Такими действиями могут быть неявка заемщика в суд, подача встречного иска по оспариванию основных параметров заключенного ранее кредитного договора (ставка по кредиту, комиссия при выдаче и др.), требование по пересмотру стоимости объекта недвижимости по текущей рыночной цене и др. Таким образом, грамотно построенная защита заемщика может существенно увеличить срок взыскания, хотя в большинстве случаев неотвратимость реализации недвижимости по той или иной цене предопределена существующим законодательством.

Заметны различия и в распределении характеристик кредита. Среди «проблемных» чаще встречались кредиты с высокой процентной ставкой и большим коэффициентом обеспечения LTV и реже — с высоким отношением плановых денежных выплат к стоимости предмета ипотеки. Эти различия не столь яркие, как расхождение в распределении заемщиков

по занятости: доли наблюдений в выделяемых по характеристикам кредита категориях различаются на 6–9 процентных пунктов в двух выборках.

Различия в распределении характеристик предмета ипотеки невелики. Среди проблемного имущества чаще встречаются дома и таунхаусы и реже — квартиры, но разность долей в двух выборках не превышает 3 процентных пункта.

5.3. Модели с объясняющими переменными

С целью моделирования связи функции риска с объясняющими переменными тестировались различные модели с наблюдаемой разнородностью: модели пропорциональных рисков (Кокса, Вейбулла, Гомперца) и ускоренного времени (логнормальная, логлогистическая и гамма-регрессия). В результате проверки правильности спецификации апробированных моделей на основании графического анализа остатков Кокса–Снелла ни одну из них, кроме непараметрической регрессии Кокса, нельзя было назвать корректной.

Для улучшения качества подгонки были использованы модели с ненаблюдаемой разнородностью, описываемой случайной величиной с гамма-распределением. Введение дополнительного параметра позволило существенно улучшить подгонку параметрических моделей. Как видно из рис. 2, проверка логнормальной и вейбулловской регрессии с помощью остатков Кокса–Снелла не обнаруживает ошибки спецификации: график оцененной интегральной функции риска для остатков выстраивается вдоль биссектрисы. Что касается заметного отклонения на графике остатков для регрессии Кокса, то оно касается правого хвоста распределения, где оценки интегрального риска имеют низкую точность. В остальной области полупараметрическая модель дает вполне приемлемое описание данных.

В таблице П2 Приложения 2 приводятся оценки параметров регрессии Кокса и логнормальной модели с ненаблюдаемым эффектом. В результате оценивания почти все факторы оказались значимыми. Наибольшее влияние на риск реализации залога оказывают факторы: тип жилья, находящегося под залогом дефолтного кредита, расположение жилья в развитом регионе и доля заемных средств в стоимости жилья.

Об интерпретации параметров уже говорилось в предыдущем разделе: потенцированные коэффициенты в модели Кокса отражают отношение функций риска между категорией кредитов, выделяемой соответствующей объясняемой переменной, и базовой категорией. Так, риск реализации квартир и домов составляет $\exp(-0.42) \times 100\% = 66\%$ от риска реализации комнат (при прочих равных условиях), т. е. комнаты реализуются значительно быстрее, чем квартиры и дома. Аналогично, если предмет ипотеки, находящийся под залогом дефолтного кредита, расположен в развитом регионе, то риск его реализации в 1.4 раза меньше, чем в других регионах. С увеличением доли заемных средств в стоимости предмета ипотеки, находящегося под залогом, риск реализации жилья увеличивается. Отношение рисков между крайними группами ($LTV \leq 50\%$ и $LTV > 70\%$) составило $\exp(0.2999 + 0.2870) \approx 1.8$. Согласно полученным оценкам, именно показатель LTV дает наибольший вклад в риск реализации.

Интересен контраст с результатом анализа описательных статистик, показавшим, что основной индикатор проблемного кредита — принадлежность заемщика к базовой категории занятости. В модели Кокса переменные занятости оказываются незначимыми, а из оценок логнормальной регрессии следует, что имущество наемных работников реализуется

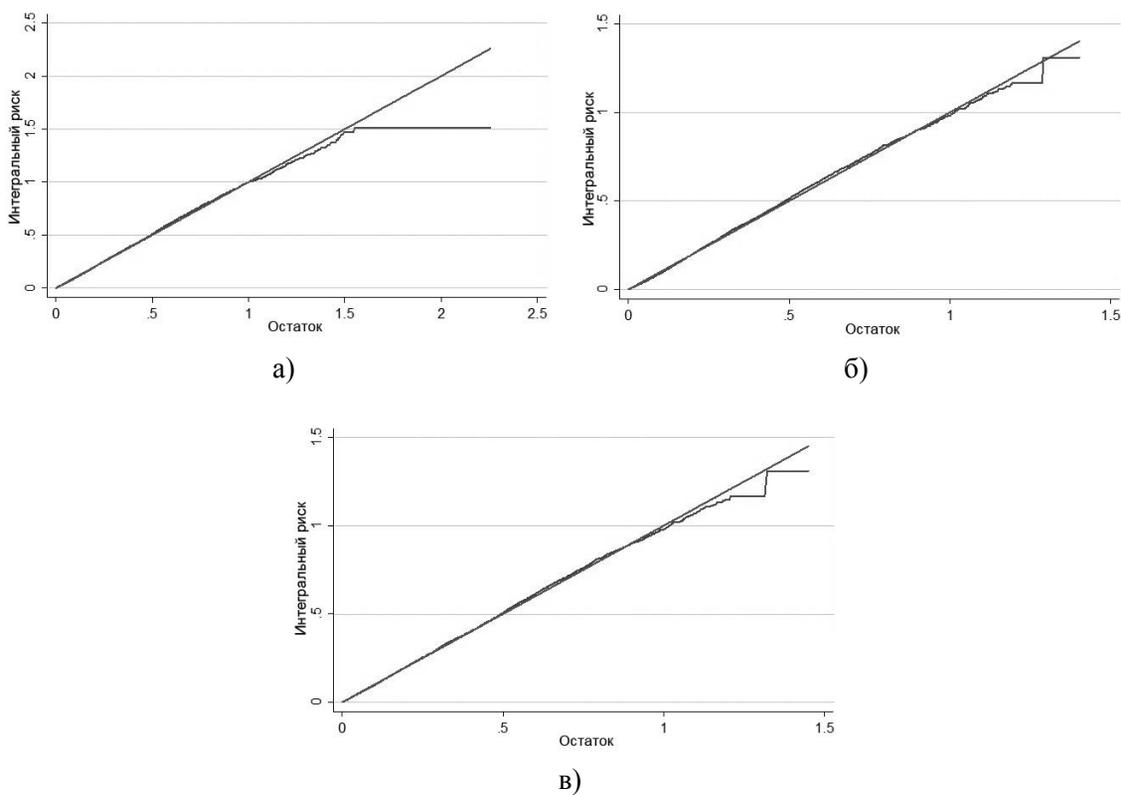


Рис. 2. Оценки интегральной функции риска для остатков Кокса–Снелла:
 а) модель Кокса; б) логнормальная регрессия с ненаблюдаемой разнородностью;
 в) регрессия Вейбулла с ненаблюдаемой разнородностью.

быстрее, чем у представителей базовой категории, но соответствующий коэффициент невелик по сравнению с коэффициентами при основных регрессорах³. Для объяснения этого контраста вновь обратимся к анализу описательных статистик.

В Приложении 3 приведено распределение объясняющих переменных в зависимости от типа занятости заемщика. Видно, что в базовой категории наблюдений намного чаще встречаются «рисковые» кредиты. Так, доля кредитов с высоким значением LTV в базовой категории составляет почти 70% против 26.4% в категории военнослужащих и госслужащих и 30.2% среди остальных наемных работников. Точно так же в базовой категории гораздо больше распространены кредиты с высокой ставкой процента и неликвидным залоговым имуществом (дом). Наконец, заемщики из базовой категории чаще встречались в наиболее развитых регионах. Из сравнения описательных статистик и оценок регрессионных моделей можно сделать следующий вывод: хотя среди заемщиков «проблемных» кредитов велика доля представителей базовой категории занятости, это связано не столько с неопределенным статусом занятости, сколько с прочими характеристиками кредита, типичными

³ В логнормальной модели положительный коэффициент отражает положительную связь объясняющей переменной со временем реализации и отрицательную — с риском. Поэтому оценки логнормальной модели и модели Кокса имеют противоположные знаки.

для заемщиков такого рода. Поэтому при включении этих характеристик в модель роль занятости становится несущественной.

В оцененных регрессионных моделях коэффициент при переменной *employ_2* (принадлежность к военнослужащим и госслужащим) оказался велик по абсолютной величине, однако этот коэффициент незначим: небольшое число наблюдений за этой категорией заемщиков не позволяет сделать надежный вывод.

5.4. Модели разделенной совокупности

Оценивались модели «mover–stayer», в которых длительность состояния в классе «mover» описывалась логнормальным распределением (далее используется обозначение Lognormal-SP+) или обобщенным гамма-распределением (Gamma-SP+), а вероятность принадлежности к классу кредитов, по которым наступает реализация — моделью логит. В логит-часть добавляется объясняющий фактор *realty_flat* — индикатор того, что жилье под залогом является квартирой. Выбор указанных опорных распределений в уравнении длительности обусловлен тем, что в простых моделях данные спецификации демонстрируют приемлемое качество подгонки. Графический анализ остатков Кокса–Снелла обеих оцененных моделей не выявляет ошибки спецификации (рис. 3).

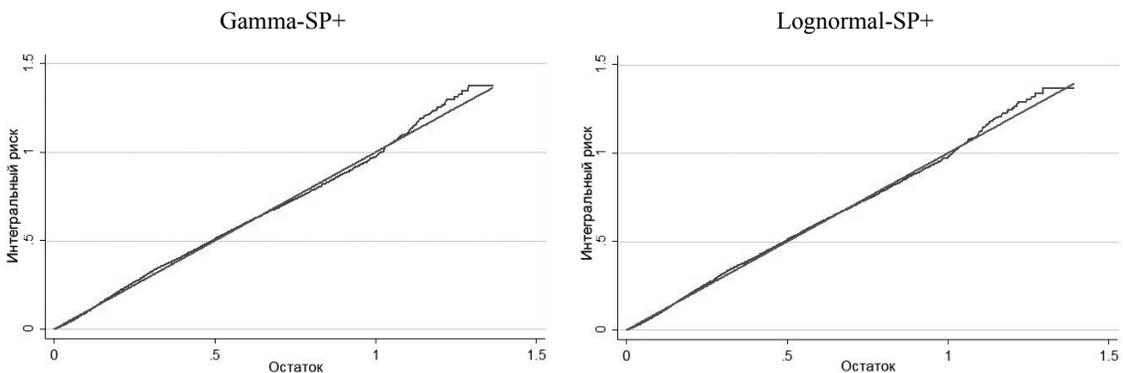


Рис. 3. Остатки Кокса–Снелла в моделях разделенной совокупности

Оценки коэффициентов модели Lognormal-SP+ приводятся в Приложении 2. В основном, они совпадают с результатами, полученными из модели Кокса и логнормальной регрессии с ненаблюдаемой разнородностью. Дополнительное уравнение позволяет оценить вероятность принадлежности кредита к классу «безнадежных» (для которых реализация, согласно модели, вообще не наступает). Оценка этой вероятности равна 32% для квартир и 24% для прочего имущества.

Качество подгонки моделей разделенной совокупности оказалось ниже, чем у моделей со случайным эффектом (для сравнения использовался критерий Акаике). Точность описания данных можно было бы увеличить за счет введения дополнительных объясняющих переменных в классифицирующее уравнение, но это затруднило бы интерпретацию результатов.

6. Заключение

В настоящей работе исследуется время жизни дефолтного кредита с первого дня просрочки до момента реализации объекта недвижимости, находящегося под залогом, в рынок в зависимости от факторов, свойственных определенной группе кредитов.

Полученные оценки дополняют результаты исследований рисков ипотечного кредитования в России (Лозинская, Ожегов, 2014; Карминский и др., 2015; Румянцева, Фурманов, 2016), давая представление о временном аспекте реализации заложенного имущества и являясь индикаторы высокого риска нереализуемости.

Проведенный анализ позволил выявить факторы, влияющие на скорость реализации имущества. Признаки, в наибольшей степени связанные с темпом реализации жилья, являются параметрами самого предмета ипотеки и кредита. Это — тип жилого помещения (дом/квартира/комната), расположение жилья в развитом регионе, а также показатель LTV, характеризующий долю заемных средств в стоимости жилья. Эти же признаки являются определяющими при анализе другого типа рискового события — дефолта по кредиту, что, однако, не является удивительным, поскольку риск реализации залога непосредственно связан с риском наступления дефолта и является по отношению к нему вторичным.

Благодарности. Авторы выражают благодарность Н. А. Фоминой за помощь в исследовании.

Список литературы

Карминский А. М., Лозинская А. М., Ожегов Е. М. (2015). Методы оценки потерь кредитора при ипотечном жилищном кредитовании. *Экономический журнал Высшей школы экономики*, 20 (1), 9–51.

Козлов А., Гушин К. (2017). Применение машинного обучения и анализа выживаемости в оценке розничных рисков. *Банковское кредитование*, 2 (72), 82–92.

Лейбов В. (2016). Как скоро банки реализуют полученное в кризис залоговое имущество. <http://bankir.ru/publikacii/20160425/kak-skoro-banki-realizuyut-poluchennoe-v-krizis-zalogovoe-imushchestvo-10007474>.

Лозинская А. М., Ожегов Е. М. (2014). Оценка кредитного риска при ипотечном жилищном кредитовании. *Прикладная эконометрика*, 35 (3), 3–17.

Ратникова Т. А., Фурманов К. К. (2014). *Анализ панельных данных и данных о длительности состояний*. М.: Издательский дом Высшей школы экономики.

Румянцева Е. В., Фурманов К. К. (2016). Моделирование времени жизни ипотечного кредита. *Прикладная эконометрика*, 41 (1), 123–143.

Aalen O. O. (1978). Nonparametric inference for a family of counting processes. *Annals of Statistics*, 6 (4), 701–726.

Altman E., Resti A., Sironi A. (2004). Default recovery rates in credit risk modelling: A review of the literature and empirical evidence. *Economic Notes by Banca dei Paschi di Siena SpA*, 33 (2), 183–208.

BCBS (2006). Basel committee on banking supervision. International convergence of capital measurement and capital standards. A revised framework comprehensive version. Bank for International Settlements. <https://www.bis.org/publ/bcbs128.htm>.

Cleves M. A., Gould W. W., Gutierrez R. G. (2004). *An introduction to survival analysis using Stata*. Revised Edition. Stata Press.

Cox D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34 (2), 187–220.

Gupton G. M. (2005). Advancing loss given default prediction models: How the quiet have quickened. *Economic Notes by Banca dei Paschi di Siena SpA*, 34 (2), 185–230.

Huang X., Oosterlee C. W. (2008). Generalized beta regression models for random loss-given-default. *Journal of Credit Risk*, 7 (4), 45–70.

Kaplan E. L., Meier P. (1958). Nonparametric estimation from incomplete observations. *Journal of American Statistical Association*, 53 (282), 457–481.

Klein J. P., Moeschberger M. L. (2005). *Survival analysis: Techniques for censored and truncated data*. 2nd Edition. N.Y.: Springer.

Kiefer N. M. (1988). Economic duration data and hazard functions. *Journal of Economic Literature*, 26 (2), 646–679.

Lozinskaya A. M. (2015). Sample selection bias in mortgage credit risk modeling. In: *Financial Econometrics and Empirical Market Microstructure* (eds. S. Ivliev, A. K. Bera, F. Lillo), 249–262. N.Y.: Springer.

Matuszyk A., Mues C., Thomas L. C. (2010). Modelling LGD for unsecured personal loans: Decision tree approach. *The Journal of the Operational Research Society*, 61 (3), 393–398.

Nelson W. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics*, 14 (4), 945–966.

Schmidt P., Witte A. D. (1987) Predicting criminal recidivism using «split population» survival time models. *NBER Working Paper* No. 2445.

Schuermann T. (2004). *What do we know about loss given default. Credit risk models and management*. 2nd Edition. London: Risk Books.

Witzany J., Rychnovský M., Charamza P. (2012). Survival analysis in LGD Modeling. *European Financial and Accounting Journal*, 7 (1), 6–27.

Zhang J., Thomas L. C. (2012). Comparisons of linear regression and survival analysis using single and mixture distributions approaches in modelling LGD. *International Journal of Forecasting*, 28 (1), 204–215.

Поступила в редакцию 25.04.2017;
принята в печать 25.08.2017.

Приложение 1

Таблица П1. Переменные, используемые в исследовании риска реализации предмета ипотеки, и их описательные статистики

Тип переменных	Переменные	Доля среди всех наблюдений, %	Доля среди проблемных* кредитов, %
Портрет основного заемщика	<i>marital</i> — 1, если основной заемщик женат/замужем, 0 — иначе	49.5	50.7
	<i>educ_level</i> — 1, если основной заемщик имеет высшее образование или ученую степень, 0 — иначе	54.4	62.7
	<i>employm_1</i> — 1, если заемщик является наемным работником, 0 — иначе	33.1	11.7
	<i>employm_2</i> — 1, если заемщик является госслужащим или военным, 0 — иначе	1.0	0.3
	<i>mainborr_contr</i> — 1, если доля дохода основного заемщика в доходе созаемщиков не более 0.8, 0 — иначе	44.1	45.2
Параметры кредита	<i>rate_high</i> — 1, если годовая кредитная ставка не менее 13.6%, 0 — иначе	40.4	48.3
	<i>LTV_50</i> — 1, если коэффициент LTV (loan to value) не превосходит 50%, 0 — иначе	9.9	11.1
	<i>LTV_70</i> — 1, если коэффициент LTV более 70%, 0 — иначе	55.8	64.7
	<i>payed_sum</i> — 1, если значение показателя** не более 1.82, 0 — иначе	37.1	31.8
Параметры предмета ипотеки	<i>floor_num_5</i> — 1, если этажность жилья не превышает 5 этажей, 0 — иначе	63.7	62.9
	<i>region_1</i> — 1, если предмет ипотеки находится в развитых регионах***, 0 — иначе	16.6	18.2
	<i>realty_flat</i> — 1, если предмет ипотеки — квартира, 0 — иначе	84.6	81.9
	<i>realty_house</i> — 1, если предмет ипотеки — дом или таунхаус, 0 — иначе	12.8	15.6

Примечание. Число наблюдений — 10681.

* — кредит относился к «проблемным», если известно, что заложенную недвижимость не удалось реализовать (или поставить на баланс банка) в течение медианного времени реализации (1909 дней).

** — формула расчета показателя *payed_sum* приведена в разделе 3.

*** — топ-10 регионов согласно рейтингу социально-экономического положения субъектов РФ по итогам 2013 года рейтингового агентства «РИА Рейтинг».

Приложение 2

Таблица П2. Оценки коэффициентов в моделях риска реализации предмета ипотеки: Кокса, логнормальной с ненаблюдаемым эффектом и Lognormal-SP+

Переменные	Регрессия Кокса	Логнормальная регрессия с ненаблюдаемым эффектом	Логнормальная регрессия с дополнительным логит-уравнением
			<i>duration</i>
<i>floor_num_5</i>	0.070* (0.034)	-0.041** (0.015)	-0.041** (0.015)
<i>rate_high</i>	-0.100** (0.033)	0.060* (0.027)	0.064** (0.024)
<i>payed_sum</i>	-0.153** (0.050)	0.041 (0.022)	0.049* (0.021)
<i>mainborr_contr</i>	0.123*** (0.031)	-0.055*** (0.015)	-0.053*** (0.014)
<i>educ_level</i>	-0.093** (0.032)	0.072*** (0.014)	0.069*** (0.014)
<i>region_1</i>	-0.344*** (0.044)	0.230*** (0.019)	0.227*** (0.020)
<i>employ_1</i>	0.083 (0.042)	-0.066*** (0.019)	-0.061*** (0.018)
<i>employ_2</i>	-0.433 (0.2439)	0.119 (0.085)	0.129 (0.084)
<i>marital</i>	-0.102** (0.031)	0.028 (0.014)	0.029* (0.014)
<i>realty_flat</i>	-0.428*** (0.077)	0.187*** (0.039)	0.164*** (0.041)
<i>realty_house</i>	-0.422*** (0.085)	0.228*** (0.043)	0.254*** (0.044)
<i>LTV_50</i>	-0.287*** (0.073)	0.070* (0.030)	0.072** (0.026)
<i>LTV_70</i>	0.300*** (0.049)	-0.125*** (0.022)	-0.115*** (0.021)
<i>_cons</i>		7.170*** (0.047)	7.118*** (0.048)
			<i>logit</i>
<i>realty_flat</i>			-0.406*** (0.098)
<i>_cons</i>			1.140*** (0.092)
	Параметры опорного распределения		
<i>ln_sig</i>		-0.832*** (0.020)	-0.794*** (0.005)
<i>ln_the</i>		0.400*** (0.056)	

Оценки каждой модели значимы в целом на уровне 0.1%.

Примечание. *, **, *** — значимость на уровне 5, 1 и 0.1% соответственно. В скобках под оценками коэффициентов — стандартные ошибки.

duration — уравнение длительности реализации.

logit — уравнение вероятности принадлежности к классу «кочевых» (для которых реализация наступает).

Приложение 3

Таблица ПЗ. Распределение переменных по типу занятости заемщика

Переменная	Наемные работники кроме госслужащих и военнослужащих	Госслужащие и военнослужащие	Остальные
<i>rate_high</i>	29.5%	24.5%	46.2%
<i>payed_sum</i>	53.4%	61.8%	28.4%
<i>floor_num_5</i>	62.4%	70.0%	64.2%
<i>mainborr_contr</i>	38.5%	64.5%	46.5%
<i>educ_level</i>	36.8%	60.0%	63.2%
<i>region_1</i>	20.1%	7.3%	35.6%
<i>marital</i>	47.9%	61.8%	50.0%
<i>realty_flat</i>	90.1%	93.6%	81.3%
<i>realty_house</i>	9.1%	6.4%	14.8%
<i>LTV_50</i>	9.2%	19.1%	10.2%
<i>LTV_70</i>	30.2%	26.4%	69.1%
Число наблюдений	3538	110	7033

Rumyantseva E. V., Furmanov K. K. Realisation of mortgage property: Survival analysis. *Applied Econometrics*, 2017, v. 48, pp. 22–43.

Ekaterina Rumyantseva

Agency for Housing Mortgage Lending, Moscow, Russian Federation; Evrumyantseva.2006@yandex.ru

Kirill Furmanov

National Research University Higher School of Economics, Moscow, Russian Federation; furmach@rbcmail.ru

Realisation of mortgage property: Survival analysis

Time to realisation of mortgage property is studied using data from Russian mortgage agency. Factors indicating high risk of non-realisation are revealed. Obtained estimates indicate that time to realisation is determined mainly by a loan-to-value ratio, a type of mortgage property and its location in economically developed region.

Keywords: mortgage; property realisation; survival models.

JEL classification: C41; G21.

References

Karminsky A., Lozinskaya A., Ozhegov E. (2015). Estimation methods of creditor's loss in residential mortgage lending. *HSE Economic Journal*, 20 (1), 9–51 (in Russian).

Kozlov A., Gushhin K. (2017). Primenenie mashinnogo obucheniya i analiza vyzhivaemosti v ocenke roznicnykh riskov. *Bankovskoe kreditovanie*, 2 (72), 82–92 (in Russian).

Lejbov V. (2016). Kak skoro banki realizujut poluchennoe v krizis zalogovoe imushchestvo. <http://bankir.ru/publikacii/20160425/kak-skoro-banki-realizuyut-poluchennoe-v-krizis-zalogovoe-imushchestvo-10007474>. (in Russian).

Lozinskaya A., Ozhegov E. (2014). Estimation of mortgage lending credit risk. *Applied Econometrics*, 35, 3–17 (in Russian).

Ratnikova T. A., Furmanov K. K. (2014). *Analiz panel'nyh dannyh i dannyh o dlitel'nosti sostojanij*. М.: Izdatel'skij dom Vysšej shkoly jekonomiki (in Russian).

Rumyantseva E., Furmanov K. (2016). Modeling mortgage survival. *Applied Econometrics*, 41, 123–143 (in Russian).

Aalen O. O. (1978). Nonparametric inference for a family of counting processes. *Annals of Statistics*, 6 (4), 701–726.

Altman E., Resti A., Sironi A. (2004). Default recovery rates in credit risk modelling: A review of the literature and empirical evidence. *Economic Notes by Banca dei Paschi di Siena SpA*, 33 (2), 183–208.

BCBS (2006). Basel committee on banking supervision. International convergence of capital measurement and capital standards. A revised framework comprehensive version. Bank for International Settlements. <https://www.bis.org/publ/bcbs128.htm>.

Cleves M. A., Gould W. W., Gutierrez R. G. (2004). *An introduction to survival analysis using Stata*. Revised Edition. Stata Press.

Cox D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34 (2), 187–220.

Gupton G. M. (2005). Advancing loss given default prediction models: How the quiet have quickened. *Economic Notes by Banca dei Paschi di Siena SpA*, 34 (2), 185–230.

Huang X., Oosterlee C. W. (2008). Generalized beta regression models for random loss-given-default. *Journal of Credit Risk*, 7 (4), 45–70.

Kaplan E. L., Meier P. (1958). Nonparametric estimation from incomplete observations. *Journal of American Statistical Association*, 53 (282), 457–481.

Klein J. P., Moeschberger M. L. (2005). *Survival analysis: Techniques for censored and truncated data*. 2nd Edition. N.Y.: Springer.

Kiefer N. M. (1988). Economic duration data and hazard functions. *Journal of Economic Literature*, 26 (2), 646–679.

Lozinskaya A. M. (2015). Sample selection bias in mortgage credit risk modeling. In: *Financial Econometrics and Empirical Market Microstructure* (eds. S. Ivliev, A. K. Bera, F. Lillo), 249–262. N.Y.: Springer.

Matuszyk A., Mues C., Thomas L. C. (2010). Modelling LGD for unsecured personal loans: Decision tree approach. *The Journal of the Operational Research Society*, 61 (3), 393–398.

Nelson W. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics*, 14 (4), 945–966.

Schmidt P., Witte A. D. (1987) Predicting criminal recidivism using «split population» survival time models. *NBER Working Paper* No. 2445.

Schuermann T. (2004). *What do we know about loss given default. Credit risk models and management*. 2nd Edition. London: Risk Books.

Witzany J., Rychnovský M., Charamza P. (2012). Survival Analysis in LGD Modeling. *European Financial and Accounting Journal*, 7 (1), 6–27.

Zhang J., Thomas L. C. (2012). Comparisons of linear regression and survival analysis using single and mixture distributions approaches in modelling LGD. *International Journal of Forecasting*, 28 (1), 204–215.

Received 25.04.2017; accepted 25.08.2017.