

А. М. Лозинская, Е. М. Ожегов

Оценка кредитного риска при ипотечном жилищном кредитовании¹

В статье рассматриваются вопросы моделирования кредитного риска на российском рынке ипотечного жилищного кредитования. Используя региональные данные по ипотечному рынку, авторы апробируют предложенную структурную модель оценки кредитного риска, которая принимает во внимание наличие проблемы самоотбора и эндогенности объясняющих переменных.

Полученные результаты могут служить основой для разработки эффективных систем риск-менеджмента в кредитных организациях.

Ключевые слова: кредитный риск; ипотечное кредитование; проблема самоотбора; эндогенность.

JEL classification: C36; D12; R20.

1. Введение

В результате формирования благоприятной законодательной базы и принятия пакета федеральных законов, направленных на повышение доступности жилья для населения, стабилизацию макроэкономической ситуации и рост реальных доходов населения, в настоящее время в России быстрыми темпами развивается рынок ипотечного жилищного кредитования (Косарева и др., 2010). Об этом, в первую очередь, свидетельствует увеличение доли ипотечных жилищных кредитов, предоставленных кредитными организациями физическим лицам, которая характеризует объем рынка ипотечного жилищного кредитования, и объемы выданных ипотечных кредитов. Объем выданных ипотечных кредитов в 2012 г. достиг уровня 1.63% от ВВП, а по итогам I квартала 2013 г. — 0.88% от ВВП². Для сравнения, в развивающихся странах объем ипотечного жилищного кредитования составляет 3–7% от ВВП, а в развитых странах этот уровень достигает 50% от ВВП (Полтерович, Старков, 2007б). Достижению таких результатов во многом способствует активная государственная жилищная политика и принятые государством антикризисные меры по поддержке ипотечного кредитования.

Вместе с тем, ипотечный кризис в США 2007–2009 гг., спровоцированный масштабными неплатежами по высокорискованным субстандартным ипотечным кредитам (sub-prime mortgages), обратил внимание на несовершенство сложившейся практики управления кредитными рисками. Кредитный риск составляет наибольшую долю совокупных рисков кредитных организаций и во многом определяет требования к размеру активов, взвешенных по уровню риска, и величине резервов на возможные потери по ссудам, а следовательно,

¹ Исследование осуществлено в рамках программы «Научный фонд НИУ ВШЭ» в 2013–2014 гг., проект № 12-01-0130.

² По данным Банка России.

и к достаточности собственного капитала. Именно от качества оценки и управления кредитным риском во многом зависит финансовое положение и жизнеспособность как отдельно взятой кредитной организации, так и всей банковской системы в целом. Поэтому центральное место в риск-менеджменте кредитной организации занимает проблема оценки вероятности дефолта (probability of default, PD) как наиболее неблагоприятной реализации кредитного риска.

2. Краткий обзор литературы

Традиционные модели оценки кредитного риска при ипотечном жилищном кредитовании используют параметрический подход к построению уравнения регрессии, описывающей вероятность дефолта (PD). В качестве таких моделей используют классические модели бинарного выбора — пробит- и логит-модели. Основная проблема, возникающая в таких моделях, известна как проблема смещения ввиду выборочной селективности или проблема самоотбора (sample selection bias problem). В результате полученные оценки являются смещенными и несостоятельными вследствие двух причин.

Первая причина связана со смещением вследствие одновременности (simultaneity bias). Такая ситуация возникает, когда при моделировании PD не принимается во внимание процесс кредитного андеррайтинга, т. е. вероятность одобрения или отклонения кредитной заявки. Другая причина связана с усечением или частичной наблюдаемостью (truncation or partial observability), которая возникает в случае отсутствия информации по отклоненным заявкам. Таким образом, величина смещения зависит от того, насколько скоррелированы процесс дефолта и процесс кредитного андеррайтинга, а также насколько полная информация по характеристикам ипотечного займа и заемщика, включая кредитную историю, содержится в выборке дефолтеров или не дефолтеров (Ross, 2000).

Определение ипотечного займа как последовательности решений заемщика и кредитной организации впервые было предложено в статье Follain (1990). Автор рассматривает процесс получения ипотечного кредита не только как оценку соотношения суммы займа к стоимости приобретаемого жилья, известного как показатель LTV (loan-to-value ratio), но и дальнейший процесс обслуживания ипотечных обязательств, в частности, обслуживание согласно условиям кредитного договора, рефинансирование ссуды или объявление дефолта, включая выбор соответствующего момента времени, а также выбор условий контракта. Rachlis, Yezer (1993) предложили теоретическую модель процесса ипотечного кредитования, включающую систему одновременных уравнений, описывающих вероятность подачи кредитной заявки, вероятность выбора конкретных параметров ипотечного займа, вероятность одобрения или отклонения кредитной заявки и вероятность дефолта.

С середины 1990-х гг. стали появляться публично доступные данные по американскому рынку ипотеки — Federal Housing Authority (FHA) foreclosure data, The Boston Fed Study data, The Home Mortgage Disclosure Act data (HMDA), а также эмпирические работы, посвященные изучению процессов принятия решений на рынке ипотечного кредитования.

Yezer et al. (1994) представили эмпирические результаты оценивания системы одновременных уравнений, включающей вероятность одобрения/отклонения кредитной заявки, сумму ипотечного займа и вероятность дефолта заемщика. Авторы, используя метод Монте-Карло, показали, что изолированное моделирование процесса одобрения/отклонения кредитной

заявки или дефолта приводит к смещенным оценкам. Позднее аналогичные результаты были получены в эмпирических работах (Munnell et al., 1996) и (Phillips, Yezer, 1996).

Ross (2000) исследовал связь между решениями об одобрении заемщика и его дефолтом. В статье (Ross, 2000) показано, что большинство оценок параметров в уравнении одобрения заемщика имеют противоположный знак в уравнении дефолта, однако это верно только после коррекции оценок на смещение выборки. В данном исследовании отмечался и тот факт, что если выборка дефолтеров / не дефолтеров содержит небольшой объем информации о характеристиках заемщика, то оцененная модель вероятности дефолта, а также модели самотбора будут страдать от существенного смещения. Чем больше доступно информации о характеристиках заемщика, включая его кредитную историю и другие факторы риска, тем выше вероятность минимизации проблемы смещения выборки.

Более поздние работы исследуют зависимость решения о выборе величины ссуды, а также других эндогенных переменных от экзогенных. Так, в работе (Ambrose et al., 2004) подчеркивается эндогенность таких факторов, как сумма ипотечного займа и коэффициент LTV. Авторы предлагают систему одновременных уравнений, включающую уравнение для описания LTV и стоимости предмета ипотечного займа, которая используется как прокси-переменная для суммы ипотечного займа при учете эндогенности.

Среди основных детерминант дефолта при ипотечном жилищном кредитовании обычно выделяют социально-демографические характеристики заемщика и параметры ипотечного займа. Примечательно, что при отсутствии данных по социально-демографическим характеристикам на индивидуальном уровне, в работе (Vajari et al., 2008) в качестве прокси-показателей используются данные по демографическим характеристикам и уровни безработицы по округам.

В исследовании (Attanazio et al., 2008) используется подход, предложенный Das et al. (2003) для оценки параметров спроса на кредиты при наличии выборочной селективности, содержащий три последовательных шага. На первом шаге оценивается вероятность наличия кредита. Второй шаг заключается в оценке уравнений для эндогенных переменных с учетом выборочной селективности. На третьем шаге оценивается уравнение спроса, где эндогенные переменные заменяются на их оцененные значения, а также присутствует компонента, отвечающая за смещение при выборочной селективности. На каждом шаге строится уравнение полиномиальной регрессии, поскольку авторы предполагают нелинейную зависимость ошибок между уравнениями.

Среди российских исследований следует отметить работу (Полтерович, Старков, 2007а), в которой обсуждаются вопросы формирования ипотечного рынка в России и даются модельные расчеты по стратегии создания массовой ипотеки в России, основанной на трансплантации модифицированного института стройсберкасс. Проблемы оценки рисков просрочки выплат по потребительским кредитам физических лиц на основе опросных данных о кредитном поведении за 2011 г. в рамках проекта «Родители и дети, мужчины и женщины в семье и обществе» рассматриваются в (Ниворожкина и др., 2013). Авторы отмечают статистически значимое влияние социально-демографических характеристик и типа региона на вероятность просрочки по выплате кредитов.

Отсутствие (закрытость) информации обуславливает ограниченность исследований российского рынка ипотечного кредитования. Основная часть эмпирических работ по моделированию дефолта посвящена американскому ипотечному рынку. В связи с этим возникает необходимость не только разработки структурной модели принятия решений на рынке ипотечного кредитования, но и ее апробации с учетом российской специфики.

В данном исследовании учтена одновременность решений заемщика с помощью использования системы одновременных уравнений. Поскольку данные для исследования являются данными одной кредитной организации, то учтена выборочная селективность данных не только на этапе кредитного андеррайтинга, но и на этапе подачи заявки в отдельно взятую кредитную организацию, а также возможности заемщика отказаться от заключения контракта после одобрения.

3. Методология исследования и данные

Структурная модель принятия решений на ипотечном рынке может быть представлена следующими шагами.

1. *Подача заявки заемщиком в кредитную организацию.* Заемщик осознает потребность в ипотеке, выбирает ипотечную программу, предлагаемую кредитными организациями, и кредитную организацию, которая в некотором смысле соответствует его предпочтениям. Потенциальный заемщик заполняет анкету в выбранной кредитной организации.

2. *Анализируя анкету заемщика, кредитная организация принимает решение о его кредитоспособности, выносит решение об одобрении/неодобрении заемщика, уточняет поданные заемщиком сведения.* В процессе кредитного андеррайтинга и в случае одобрения кредитные организации в некоторых случаях также устанавливают предельную величину ссуды для данного заемщика.

3. *Одобренный заемщик принимает решение о заключении кредитного контракта или не заключении такового.* В случае заключения контракта он выбирает величину ссуды (не выше предельной), срок погашения, первоначальный взнос и размер ежемесячного платежа в соответствии с условиями, определенными кредитной программой, а также объект недвижимости для приобретения.

4. *Заемщик выбирает стратегию обслуживания долга:* обслуживать контракт в соответствии с условиями либо погасить кредит досрочно, допустить просрочку, объявить дефолт или рефинансировать долг.

Эконометрическая модель повторяет шаги структурной модели.

На первом шаге моделируется вероятность подачи заявки на получение ипотечного жилищного кредита y_{1it} , для оценивания которой используются агрегированные данные по региональному рынку ипотеки:

$$y_{1it} = g_{1t}(z_{1t}) + e_{1it}, \quad (1)$$

$$e_{1it} \sim \text{iid}(0, \sigma_{e_{1t}}^2), z_{1t} = M_t,$$

где $t = 1, \dots, T$; T — число месяцев; y_{1it} — вероятность подачи заявки i -м потенциальным заемщиком в момент времени t (рассчитывается как число поданных заявок в месяц, деленное на число домашних хозяйств), одинаковая для заемщиков, подающих заявку в один месяц; M_t — вектор экзогенных макроэкономических показателей.

На втором шаге оценивается уравнение вероятности одобрения кредитной заявки y_{2it}^* для всех поданных заявок:

$$y_{2it}^* = g_2(z_{2it}^*) + e_{2it}, \tag{2}$$

$$y_{2it} = y_{1it} y_{2it}^* \text{ — наблюдаемо,}$$

$$\text{cov}(e_{1it}, e_{2it}) = \sigma_{12}, \quad i \in N_t, \quad \text{cov}(e_{1it}, e_{2it}) = 0, \quad i \notin N_t, \quad z_{2it}^* = (D_i^*, M_t),$$

где N_t — набор индивидов, подавших заявку в момент времени t ; $N = (N_1, \dots, N_T)$; y_{2it} — решение об одобрении кредитной заявки i -го заемщика; z_{2it}^* — вектор экзогенных индивидуальных демографических характеристик и макроэкономических показателей на дату подачи кредитной заявки.

В случае одобрения кредитной заявки кредитная организация выставляет кредитный лимит. В случае одобрения без выставления лимита данный шаг и использование данной переменной можно опустить. Кредитный лимит является эндогенной переменной, т. к. выставляется кредитной организацией с учетом потенциальных решений заемщика на следующих шагах. Стандартные методы оценивания в таком случае будут давать смещенные и несостоятельные оценки. Для преодоления данной проблемы используется метод инструментальных переменных, где далее в качестве инструментов используются прогнозные значения эндогенных переменных:

$$\bar{L}_{it}^* = g_{\bar{L}}(z_{\bar{L}it}^*) + e_{\bar{L}it}, \tag{3}$$

$$\bar{L}_{it} = y_{2it} \bar{L}_{it}^* \text{ — наблюдаемо, } \text{cov}(e_{2it}, e_{\bar{L}it}) = \sigma_{2\bar{L}},$$

где \bar{L} — величина кредитного лимита; $z_{\bar{L}it}^*$ — вектор инструментальных демографических характеристик и макроэкономических показателей на дату подачи кредитной заявки.

Следующий этап включает в себя моделирование вероятности заключения кредитного контракта потенциальным заемщиком:

$$y_{3it}^* = g_3(z_{3it}^*, \hat{L}_{it}^*) + e_{3it}, \tag{4}$$

$$y_{3it} = y_{2it} y_{3it}^* \text{ — наблюдаемо, } \text{cov}(e_{2it}, e_{3it}) = \sigma_{23}, \quad z_{3it}^* = (D_i^*, M_t), \quad \hat{L}_{it}^* = \hat{g}_{\bar{L}}(z_{\bar{L}it}^*),$$

где $y_{3it} = 1$ в случае заключения кредитного контракта; z_{3it}^* — вектор индивидуальных демографических характеристик и макроэкономических показателей на дату подачи кредитной заявки; \hat{L}_{it}^* — прогнозная величина кредитного лимита.

Одновременный выбор параметров ипотечного займа и стоимости приобретаемого жилья для всех заемщиков, готовых для заключения кредитного контракта, описывается следующим образом:

$$\begin{cases} C_{jit}^* = g_{C_j}(z_{C_{jit}^*}, \hat{L}_{it}^*, C_{-jit}^*, V_i^*) + e_{ji}, \\ V_{it}^* = g_V(z_{V_{it}^*}, \hat{L}_{it}^*, C_{it}^*) + e_{Vi}, \end{cases} \tag{5}$$

$$(C_{it}, V_{it}) = y_{3it} (C_{it}^*, V_{it}^*) \text{ — наблюдаемо,}$$

$$\text{cov}(e_{3it}, e_{ji}) = \sigma_{3j}, \quad \text{cov}(e_{3it}, e_{Vi}) = \sigma_{3V}, \quad \text{cov}(e_{ji}, e_{ki}) = \sigma_{jk}, \quad \text{cov}(e_{ji}, e_{Vi}) = \sigma_{jV},$$

$$z_{it}^* = (D_i^*, M_t, F_i),$$

где V — стоимость приобретаемого жилья; $C = (C_j, C_{-j})$ — вектор параметров ипотечного займа (сумма кредита, срок кредита, размер первоначального взноса, процентная ставка по кредиту, тип процентной ставки); z_{it}^* — вектор индивидуальных демографических характеристик D_i^* , макроэкономических показателей на дату подачи кредитной заявки M_t и характеристик приобретаемого жилья F_t .

Вероятность дефолта (PD) y_{4it}^* описывается как

$$y_{4it}^* = g_4(z_{4it}^*, \hat{C}_{it}^*, \hat{V}_{it}^*) + e_{4i}, \tag{6}$$

$$y_{4it} = y_{3it} y_{4it}^* \text{ — наблюдаемо, } \text{cov}(e_{3i}, e_{4i}) = \sigma_{34},$$

где \hat{C}_{it}^* — прогнозные значения параметров кредитного контракта; \hat{V}_{it}^* — прогнозная стоимость приобретаемого жилья.

В (Das et al., 2003) показано, что с некоторыми незначительными предположениями модель выборочной селективности с эндогенными регрессорами вида

$$y_i^* = g_0(z_{1i}^*, x_i) + e_i, \tag{7}$$

$$x_i^* = \pi(z_{1i}^*, z_{2i}^*) + v_i,$$

$$y_i = dy_i^* \text{ наблюдается, } x_i = dx_i^* \text{ наблюдается,}$$

$$\text{cov}(e_i, d) \neq 0, \text{cov}(v_i, d) \neq 0$$

может быть оценена вплоть до аддитивной константы при использовании следующей трехшаговой процедуры:

- 1) состоятельная оценка $\hat{p} = E[d | z_d]$;
- 2) оценка $\hat{x} = E[x | z, d = 1] = \pi(z_1, z_2) + \mu_0(\hat{p})$;
- 3) оценка $E[y | x, z, d = 1] = g_0(z, \hat{x}) + \lambda_0(\hat{p})$.

Данная процедура может быть расширена на случай нетреугольной системы одновременных уравнений с выборочной селективностью и эндогенными регрессорами (5) при помощи добавления дополнительного шага. Так, для модели вида

$$\begin{cases} y_{1i}^* = g_1(z_{1i}^*, y_{-1,i}^*, x_i) + e_{1i}, \\ \dots \\ y_{Ki}^* = g_K(z_{Ki}^*, y_{-K,i}^*, x_i) + e_{1i}, \end{cases} \tag{8}$$

$$y^* = (y_1^*, \dots, y_K^*) = (y_j^*, y_{-j}^*), z^* = (z_1^*, \dots, z_K^*), x_i^* = \pi(z_i^*, z_{xi}^*) + v_i,$$

$$y_{ji} = dy_{ji}^* \text{ наблюдается, } j = 1, \dots, K, x_i = dx_i^* \text{ наблюдается,}$$

$$\text{cov}(e_{ji}, e_{ki}) = \sigma_{jk}, j, k \in \{1, \dots, K\},$$

$$\text{cov}(e_{ji}, d) = \sigma_{jd}, j \in \{1, \dots, K\}, \text{cov}(v_i, d) = \sigma_{vd}$$

процедура оценки будет выглядеть как:

- 1) состоятельная оценка $\hat{p} = E[d | z_d]$;
- 2) оценка $\hat{x} = E[x | z, z_x, d = 1] = \pi(z, z_x) + \mu_0(\hat{p})$;

3) оценка $\hat{y}_j = E[y_j | z, x, d = 1] = g'_j(z, \hat{x}) + \lambda_j(\hat{p})$;

4) оценка $E[y_j | z, y_{-j}, x, d = 1] = g_j(z, \hat{y}_{-j}, \hat{x}) + \lambda_j(\hat{p})$.

Эта процедура обеспечивает идентификацию g вплоть до K аддитивных констант, если μ_0, λ, π, g непрерывно дифференцируемы почти всюду, $\frac{\partial \hat{p}(z_d)}{\partial z_d} \neq 0$ почти наверное и $\text{rank} \frac{\partial \hat{x}}{\partial z_x} = \dim(x)$.

Для оценки модели (1)–(6) были сделаны следующие предположения. Ошибки в уравнениях (1)–(6) имеют совместное нормальное распределение с нулевым вектором первых моментов. Матрица центрированных вторых моментов содержит диагональные элементы, равные единице (ввиду идентифицируемости модели с точностью до набора из k констант необходимо задать их конкретные значения), и ненулевой ковариации между ошибками, которая будет оценена. Это предположение подразумевает использование лямбды Хекмана (обратного отношения Миллса) для μ_0 и λ функций в уравнениях (7)–(8) (Heckman, 1976, 1979). В сочетании с аппроксимацией функций π и g полиномами первой степени данное предположение позволяет удовлетворить условие идентифицируемости модели. В работе Attanasio et al. (2008) показано, что оценки в уравнении спроса на кредит с помощью полиномиальной аппроксимации более высокой степени также являются состоятельными, но менее эффективными. Доказательство этого вывода для имеющегося набора данных в рамках настоящего исследования проводиться не будет.

Для цели исследования были использованы два набора данных. Первый набор представляет собой агрегированные региональные ежемесячные данные, характеризующие развитие рынка ипотечных жилищных кредитов, застрахованных государством, и региональные макроэкономические показатели за период с 01.08.2008 до 31.08.2012. Второй набор содержит информацию одного из региональных представительств Агентства по ипотечному жилищному кредитованию (АИЖК) по 4300 российским заемщикам, подавшим заявку на получение ипотечного займа в период 2008–2012 гг. Среди доступных показателей следует отметить социально-демографические характеристики заемщика на дату подачи ипотечной заявки, параметры ипотечного договора, характеристики предмета ипотеки, а также информацию по обслуживанию ипотечного займа. Набор используемых переменных и их описательные статистики представлены в табл. 1 и 2 Приложения. Далее при оценивании модели (1)–(6) категориальные переменные дохода заемщика и созаемщиков, уровня образования, семейного положения и места работы были перекодированы в набор бинарных переменных.

Первоначально объем выборки составлял 4897 наблюдения. Однако далее был проведен предварительный анализ данных на предмет наличия статистических выбросов и явных ошибок измерения. В частности, из выборки были исключены:

- заемщики, чей возраст не был определен или был менее 21 года;
- ипотечные займы с отрицательным и/или нулевым ежемесячным платежом и/или процентной ставкой;
- наблюдения с показателем LTV, превышающим 1 или близким к 0;
- наблюдения с показателем DTI (соотношение ежемесячного платежа по кредиту к ежемесячному доходу заемщика), равным 0 или превышающим 1.

Примечательно, что выборка включала в себя информацию как по одобренным, так и по отклоненным кредитным заявкам (в соотношении 86:14). При этом только 76.6% одобренных клиентов приняли решение о заключении ипотечной сделки, среди которых общая доля

дефолтов, т. е. ситуации просроченных платежей на срок более чем 90 дней, составляет 5%. Проблема несбалансированности выборки характерна для задач моделирования кредитного риска. Однако при построении моделей бинарного выбора и даже модели линейной вероятности несбалансированность оказывает влияние только на оценку свободного члена, но не на оценки коэффициентов при объясняющих переменных (Maddala, 1992, p. 325).

Параметры ипотечного кредита представляют собой прокси-переменные для оценки кредитного риска отдельного заемщика. По сути, ипотечные продукты с высоким LTV более привлекательны для неликвидных заемщиков, многие из которых будут сталкиваться с проблемами при погашении займа. Так, например, заемщики с $LTV > 0.9$ живут с «психологией арендатора», т. к. не вложили много средств в свое жилье, поэтому не очень мотивированы бороться за сохранение кредита при малейших сложностях с обслуживанием. Кредиты с высоким показателем LTV являются высокорискованными, и свои риски кредиторы компенсируют за счет высокой процентной ставки.

Как видно из табл. 1 Приложения, выборка имеет высокую вариацию LTV при среднем значении 0.56. Другими словами, в среднем заемщики являются не высокорискованными и приобретающими жилье со средней стоимостью 2 млн руб., что является довольно типичной ситуацией для вторичного рынка недвижимости региона.

Один из ключевых параметров ипотечного кредита — процентная ставка. Как правило, кредитные организации предлагают ипотечные продукты с фиксированной и плавающей процентной ставкой. Поскольку размер плавающей процентной ставки привязан к уровню одного из рыночных индикаторов, такие кредитные продукты обычно характеризуются более высоким уровнем кредитного риска. В табл. 2 Приложения показано, что только 13.5% ипотечных кредитов имеют плавающую процентную ставку.

Среди социально-демографических характеристик доход заемщика во многом определяет его способность в будущем погашать ипотечные обязательства, поэтому данный показатель играет важную роль в объяснении вероятности подачи кредитной заявки, вероятности одобрения кредитной заявки, вероятности заключения кредитного контракта и вероятности дефолта. Примечательно, что из 4300 индивидов 67.86% не отразили информацию о своем доходе, 8.74 и 13.93% индивидов в выборке имеют ежемесячный доход до 20 и до 40 тыс. руб. соответственно. В среднем, 45% ежемесячного дохода направляется на погашение ежемесячного ипотечного жилья, о чем свидетельствует показатель DTI в табл. 1 Приложения.

Уровень образования может рассматриваться в качестве прокси-показателя для уровня финансовой грамотности населения, который оказывает влияние на вероятность дефолта заемщика. Как видно из табл. 2. Приложения, основная часть клиентов имеет высшее (52.36%) и среднее образование (42.69%).

Макроэкономические показатели определяют рыночный спрос и предложение на ипотечные жилищные кредиты, застрахованные государством. Коэффициент доступности жилья в среднем составляет 3.48 года, что указывает период времени, в течение которого заемщик может накопить на квартиру при условии, что все получаемые денежные доходы будут откладываться на ее приобретение. В Европе значение данного показателя варьируется от 3 до 6 лет. Однако существует значительная разница между Россией и развитыми странами с точки зрения условий осуществления накопления для решения жилищной проблемы, а также доступа к кредитным ресурсам (Основы ипотечного кредитования, 2006).

4. Эмпирические результаты

Для выявления основных детерминант вероятности дефолта была построена пробит-модель, учитывающая корректировку на проблему самоотбора и эндогенность объясняющих переменных. Принимая во внимание, что для оценивания модели использовалась информация только по заключенным контрактам, число наблюдений составило 2600. Эмпирические результаты представлены в табл. 3 Приложения.

Параметры модели (1)–(6) были оценены с помощью линейных функций g для непрерывных объясняемых переменных, пробит g -функций для дискретных объясняемых переменных с показателем смещения (из-за выборочной селективности) в виде лямбды Хекмана. Было проведено сравнение полученных оценок с теми, которые не были скорректированы на выборочную селективность и эндогенность. Коэффициенты перед показателями смещения ввиду выборочной селективности оказались значимыми во всех уравнениях, что говорит о необходимости их включения в модель. Тест на нормальность распределения подтвердил гипотезу о том, что остатки модели после корректировки на выборочную селективность и эндогенность распределены нормально. Стандартные ошибки параметров были также оценены робастным способом (для контроля неучтенной гетерогенности заемщиков) и с помощью бутстрапа (для контроля возможной корреляции между заемщиками, которые подали заявки на ипотечный кредит в одном месяце) со 100 повторениями. По сравнению с обычными оценками стандартных ошибок параметров, оценки параметров с бутстрап-стандартными ошибками сохраняют значимость во всех спецификациях.

Оценки параметров уравнений (1)–(5), являющихся предварительными для оценки уравнения вероятности дефолта, имеют ожидаемые знаки.

На шаге 1 были получены оценки вероятности подачи кредитной заявки с помощью модели линейной вероятности, статистически значимый вклад в объяснение которой вносят макроэкономические показатели региона, характеризующие спрос и предложение на рынке ипотечных жилищных кредитов. Для данной модели были использованы 4284 наблюдения после исключения наблюдений с пропущенными переменными. Так, вероятность подачи заявки в отделение АИЖК возрастает при негативных шоках спроса (рост цен на недвижимость и снижение доходов населения) или предложения на рынке ипотеки (увеличение средних ставок, снижение средних сроков выдаваемых кредитов).

На шаге 2 для получения состоятельных оценок вероятности одобрения кредитной заявки использовалась пробит-модель с корректировкой на смещение выборки. Для данной модели были использованы 3987 наблюдений после исключения наблюдений с пропущенными переменными. Примечательно, что процесс кредитного андеррайтинга региональных отделений АИЖК во многом определяется не только характеристиками потенциальных заемщиков, но и числом поданных кредитных заявок в конкретный период времени, которые отрицательно влияют на вероятность одобрения кредитной заявки. Данный факт объясняется наличием лимитов у региональных представительств, выставляемых АИЖК. Поэтому в процессе кредитного андеррайтинга региональные отделения должны принимать во внимание наличие лимитов и регулировать число одобренных заявок.

На шаге 3 были получены состоятельные оценки вероятности заключения кредитного договора на основе оценивания пробит-модели с корректировкой на смещение выборки и эндогенность кредитного лимита. Для данной модели были использованы 3487 наблюдений после исключения наблюдений с пропущенными переменными. Статистически значимое

положительное влияние на вероятность заключения контракта оказывает семейное положение и место работы, а также сумма займа, величина выставленного кредитного лимита и срок ипотеки наряду с низкой процентной ставкой.

На последнем шаге была оценена модель вероятности дефолта заемщика как пробит-модель бинарного выбора с корректировкой на выборочную селективность ввиду неодобрения заемщиков и их отказа от заключения контракта, а также корректировки на эндогенность параметров кредитного контракта. Для данной модели были использованы 2229 наблюдений после исключения наблюдений с пропущенными переменными. Положительный статистически значимый вклад в объяснение вероятности дефолта вносят такие параметры ипотечного займа, как размер процентной ставки, срок кредита и сумма ипотечного займа. Это подтверждает гипотезу о том, что параметры ипотечного кредита служат прокси-переменными для оценки кредитного риска отдельно взятого заемщика.

Вероятность дефолта сокращается при увеличении числа созаемщиков. Принимая во внимание солидарную ответственность созаемщиков по ипотечному кредиту, основной заемщик будет соблюдать тем более высокую дисциплину платежей, чем больше у него созаемщиков, ввиду потенциально более высоких моральных издержек от неплатежей.

Среди социально-демографических характеристик следует обратить внимание на наличие параболической зависимости вероятности дефолта от возраста заемщика. Данный факт можно объяснить тем, что молодое и старое поколение более склонно к краткосрочному планированию. Поэтому при возникновении сложностей с погашением ипотечного кредита, который по своей природе является долгосрочным кредитом, данные категории населения более склонны к принятию решения об объявлении дефолта.

Статистически значимые положительные оценки параметров корреляции между ошибками в уравнениях вероятности дефолта и подачи заявки в АИЖК, а также вероятности дефолта и одобрения заемщика позволяют сделать выводы о предпочтениях банка относительно риска. Так, положительная корреляция между решением об одобрении заемщика и его вероятным дефолтом говорит, что банк не вполне заботится о том, чтобы отсеивать заемщиков по финансовым соображениям. Наличие программ для высокорискованных категорий заемщиков, а также процесс андеррайтинга свидетельствует о том, что АИЖК в первую очередь ориентирован на выполнение социальных функций по обеспечению доступным жильем заемщиков, испытывающих сложности в получении ипотечного кредита в коммерческом банке. Положительная корреляция между вероятностью подачи заявки в АИЖК и вероятностью дефолта также свидетельствует о том, что в моменты негативных шоков ипотечного рынка кредитная организация, выполняя социальную программу, выбирает стратегию установления доступных цен на ипотеку (при росте рыночных ставок и снижении общих объемов рынка увеличивается вероятность подачи заявки в АИЖК) вопреки росту рыночного риска.

5. Заключение

Причины и последствия ипотечного кризиса в США 2007–2009 гг. обусловили повышенное внимание как органов банковского надзора, так и коммерческих организаций к методам оценки кредитного риска и способам их совершенствования. Предложенная в работе структурная эконометрическая модель оценки кредитного риска позволяет моделировать процесс

ипотечного кредитования и выявлять основные факторы, определяющие вероятность подачи кредитной заявки и ее одобрения, вероятность заключения ипотечного договора и дефолта. Предложенная многошаговая процедура оценивания кредитного риска позволяет получить состоятельные оценки вероятности дефолта в результате корректировки на эндогенность и выборочную селективность.

Результаты апробирования модели на агрегированных региональных данных по ипотечному рынку и индивидуальных данных регионального представительства АИЖК указывают на наличие зависимости вероятности дефолта от социально-демографических характеристик индивида. Примечательно, что к наиболее ненадежным заемщикам относятся индивиды среднего возраста и среднего уровня дохода. Параметры ипотечного займа служат прокси-переменными для оценки кредитного риска и также вносят положительный статистически значимый вклад в объяснение вероятности дефолта.

Оценка корреляции между ошибками в уравнениях подачи заявки в АИЖК и ее одобрением дает основание утверждать, что заемщики по ипотечным жилищным кредитам, выдаваемым АИЖК, чаще, чем средние индивиды, сталкиваются с проблемами при обслуживании займа. Другими словами, они характеризуются более высоким уровнем кредитного риска. Данный факт во многом объясняется тем, что АИЖК выполняет социальные функции по обеспечению доступным жильем в том числе высокорискованных групп заемщиков.

Предложенная методика оценки кредитного риска и полученные результаты могут служить основой для надстройки действующих систем риск-менеджмента кредитных организаций, а также выступать как элемент внутренней системы оценки кредитного риска по ипотечным кредитам.

Среди ограничений исследования следует обозначить особенности используемых данных, в частности, отсутствие информации о кредитной истории клиентов, причин не заключения сделок в случае одобренных ипотечных заявок, а также низкая вариация в агрегированных демографических характеристиках. В качестве возможных направлений дальнейших исследований следует отметить выявление нелинейных взаимосвязей, использование более гибких инструментов эконометрического оценивания в классе непараметрических и полупараметрических методов, а также сравнительный анализ с классическими инструментами параметрического оценивания.

Список литературы

Косарева Н. Б., Копейкин А. Б., Рогожина Н. Н., Сиваев Д. Б., Туманов А. А. (2010). *Развитие ипотечного кредитования в Российской Федерации*. М.: Издательство «Дело» РАНХ.

Ниворожкина Л. Н., Овчарова Л. Н., Синявская Т. Г. (2013). Эконометрическое моделирование риска выплат по потребительским кредитам. *Прикладная эконометрика*, 30 (2), 65–76.

Основы ипотечного кредитования. (2006). М.: Фонд «Институт экономики города».

Полтерович В. М., Старков О. Ю. (2007а). *Формирование ипотеки в догоняющих экономиках: проблема трансплантации институтов*. М.: Наука.

Полтерович В. М., Старков О. Ю. (2007б). Стратегия формирования ипотечного рынка в России. *Экономика и математические методы*, 43 (4), 3–22.

Ambrose B., LaCour-Little M., Sanders A. (2004). The effect of conforming loan status on mortgage yield spreads: A loan level analysis. *Real Estate Economics*, 32 (4), 541–569.

Attanasio O. P., Goldberg P. K., Kyriazidou E. (2008). Credit constraints in the market for consumer durables: Evidence from micro data on car loans. *International Economic Review*, 49 (2), 401–436.

Bajari P., Chu C. S., Park M. (2008). An empirical model of subprime mortgage default from 2000 to 2007. *NBER Working Paper 14625*.

Das M., Newey W. K., Vella F. (2003). Nonparametric estimation of sample selection models. *The Review of Economic Studies*, 70 (1), 33–58.

Follain J. R. (1990). Mortgage choice. *Real Estate Economics*, 18 (2), 125–144.

Heckman J. (1976). The common structure of statistical models of truncation, sample selection, and limited dependent variables and a sample estimator for such models. *Annals of Economic and Social Measurement*, 5 (4), 475–492.

Heckman J. (1979). Sample selection bias as a specification error. *Econometrica*, 47 (1), 153–161.

Maddala G. S. (1992). *Introduction to econometrics*. 2nd ed. Macmillan.

Munnell A. H., Tootell G. M., Browne L. E., McEneaney J. (1996). Mortgage lending in Boston: Interpreting HMDA data. *American Economic Review*, 86, 25–53.

Phillips R., Yezer A. (1996). Self-selection and tests for bias and risk in mortgage lending: Can you price the mortgage if you don't know the process? *Journal of Real Estate Research*, 11, 87–102.

Rachlis M., Yezer A. (1993). Serious flaws in statistical tests for discrimination in mortgage markets. *Journal of Housing Research*, 4, 315–336.

Ross S. L. (2000). Mortgage lending, sample selection and default. *Real Estate Economics*, 8, 581–621.

Yezer A., Phillips R., Trost R. (1994). Bias in estimates of discrimination and default in mortgage lending: The effects of simultaneity and self-selection. *Journal of Real Estate Finance and Economics*, 9, 197–215.

Приложение

Таблица 1. Определение переменных и описательные статистики

Переменные	Среднее	Станд. отклонение	Min	Max
Решение об одобрении кредитной сделки: 1 — кредитная сделка одобрена 0 — в противном случае	—	—	—	—
Решение о заключении контракта: 1 — клиент заключил кредитную сделку 0 — в противном случае	—	—	—	—
Вероятность дефолта: 1 — если дефолт (просроченная задолженность больше 90 дней) 0 — в противном случае	—	—	—	—
Возраст клиента, лет	34	7.6	21	61
Квадрат возраста клиента, лет ²	—	—	—	—
Число созаемщиков, шт.	0.62	0.57	0	3
Максимальная сумма кредитного лимита, тыс. руб.	936.1	684.95	0	12700
Процентная ставка по кредиту, %	11.58	1.62	9.55	19
Сумма ипотечного займа, тыс. руб.	1040.0	573.5	120	10000

Окончание табл. 1

Переменные	Среднее	Станд. отклонение	Min	Max
Размер первоначального взноса, тыс. руб.	854.5	706.7	0	13800
Отношение суммы займа к оценочной стоимости приобретаемого жилья (LTV)	0.56	0.17	0.02	0.94
Отношение размера ежемесячного платежа к ежемесячному доходу заемщика (DTI)	0.45	0.18	0.06	1
Оценочная стоимость приобретаемого жилья, тыс. руб.	1894.5	1049.3	330	15300
Наблюдаемое число дней кредита	786.65	430.77	15	1487
Квартальный уровень безработицы в регионе, %	8.43	1.51	6.3	10.9
Средний размер ипотечного займа в регионе, тыс. руб.	1160.27	252.23	899.31	1908.2
Медианный срок ипотечного займа в регионе, мес.	201.64	12.7	173	222.2
Медианная процентная ставка по ипотечному кредиту в регионе, %	13.1	0.82	12	14.3
Среднее отношение размера ежемесячного платежа к ежемесячному доходу заемщика в регионе, %	34.81	0.7	33.44	36.68
Средняя стоимость 1 кв. м жилья в регионе, тыс. руб.	37.6	6.4	28.8	51.3
Коэффициент доступности жилья, лет	3.48	0.68	2.57	4.65
Объем выданных ипотечных кредитов в регионе, тыс. руб.	885948.4	563161.6	116100	2191000
Число выданных ипотечных кредитов в регионе, шт.	896.57	528.89	134	2112

Таблица 2. Определение и описательные статистики по категориальным переменным

Переменные	Число наблюдений	Доля, %
Пол:		
1 — мужской	1881	43.74
0 — женский	2419	56.26
всего	4300	
Семейное положение:		
1 — не замужем/холостяк	1221	28.70
2 — замужем/женат	2359	55.45
3 — вдовец/вдова	56	1.32
4 — в разводе	618	14.53
всего	4254	
Место работы:		
1 — не работает	1	0.02
2 — пенсионер ¹	0	0.00
3 — военнослужащий	13	0.31
4 — наемный работник	3965	95.27
5 — индивидуальный предприниматель	39	0.94
6 — госслужащий	144	3.46
всего	4162	

³ В выборке присутствовал один пенсионер. После предварительного анализа данных данное наблюдение было исключено.

Окончание табл. 2

Переменные	Число наблюдений	Доля, %
Уровень образования:		
1 — начальное	65	1.59
2 — среднее	1748	42.69
3 — незаконченное высшее	138	3.37
4 — высшее	2144	52.36
всего	4095	
Ежемесячный доход заемщика:		
0 — данные о доходе отсутствуют	2918	67.86
1 — 0–9999 руб.	118	2.74
2 — 10 000–19 999 руб.	376	8.74
3 — 20 000–39 999 руб.	599	13.93
4 — ≥ 40 000 руб.	289	6.72
всего	4300	
Сумма основного ежемесячного дохода созаемщиков:		
0 — данные о доходе отсутствуют	3725	86.63
1 — 0–9999 руб.	160	3.72
2 — 10 000–19 999 руб.	225	5.23
3 — ≥ 20 000 руб.	190	4.42
всего	4300	
Тип процентной ставки:		
0 — фиксированная	2 423	86.50
1 — плавающая	378	13.50
всего	2801	
Срок кредита:		
1 — < 120 мес.	181	6.46
2 — 120–179 мес.	595	21.25
3 — 180–239 мес.	1107	39.54
4 — 240–299 мес.	690	24.64
5 — ≥ 300 мес.	227	8.11
всего	2800	

Примечание. По переменным «семейное положение», «место работы» и «уровень образования» в части наблюдений данные отсутствовали. Доля рассчитана как процент от общего числа фактически имеющих наблюдения. Данные о типе процентной ставки и сроке кредита доступны только для выданных ипотечных кредитов.

Таблица 3. Результаты оценки пробит-модели для вероятности дефолта с учетом корректировки на проблему самоотбора и эндогенность объясняющих переменных

Переменные	Коэффициенты
Вероятность подачи ипотечной заявки	0.061 ^{***} (0.021)
Вероятность одобрения ипотечной заявки	1.580 [*] (0.952)
Вероятность заключения ипотечного договора	0.510 (0.642)
Плавающая процентная ставка	-11.641 (20.548)

Окончание табл. 3

Переменные	Коэффициенты
Размер процентной ставки	1.659*** (0.371)
Срок кредита < 120 месяцев	8.310 (8.659)
Срок кредита 120–179 месяцев	15.298** (7.714)
Срок кредита 180–239 месяцев	14.128* (7.263)
Срок кредита 240–299 месяцев	29.278*** (10.946)
Оценочная стоимость жилья, тыс. руб.	–0.009*** (0.003)
Сумма займа, тыс. руб.	0.018*** (0.006)
LTV	–36.063** (14.353)
Возраст заемщика	–0.742** (0.367)
Квадрат возраста заемщика	0.0106* (0.0060)
Мужской пол	0.255** (0.124)
Заемщик — государственный служащий	0.776 (0.519)
Число созаемщиков	–0.349*** (0.112)
Отсутствуют данные о ежемесячном доходе	0.392 (0.670)
Ежемесячный доход 10 000–19 999 руб.	–0.427* (0.230)
Ежемесячный доход 20 000–39 999 руб.	–0.080 (0.397)
Ежемесячный доход ≥ 40 000 руб.	–0.130 (0.659)
Средняя стоимость 1 кв. м жилья в регионе, тыс. руб.	–0.054** (0.021)
Наблюдаемое число дней кредита	0.0005 (0.0010)
Константа	–9.452 (8.161)
Число наблюдений	2600

Примечание. Базисные категории: заемщик — наемный работник, ежемесячный доход 0–9999 руб., срок кредита ≥ 300 месяцев. ***, **, * — значимость на уровне 1, 5 и 10% соответственно. В круглых скобках приведены бутстрап-стандартные отклонения коэффициентов для 100 повторений.