



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

# ОЦЕНКА ВЕРОЯТНОСТИ ДЕФОЛТА ПРИ ПРИ ИПОТЕЧНОМ ЖИЛИЩНОМ КРЕДИТОВАНИИ С ИСПОЛЬЗОВАНИЕМ СИСТЕМЫ ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ

Порошина Агата Максимовна

Ожегов Евгений Максимович

Кафедра прикладной математики и моделирования в социальных системах

Лаборатория междисциплинарных эмпирических исследований:

Сектор эмпирического анализа рынков и компаний

# Структура доклада

1. Актуальность исследования
2. Стилизованные факты
3. Предыдущие исследования
4. Исследовательские вопросы
5. Данные
6. Эконометрическая модель
7. Эмпирические результаты
8. Выводы

В данной научной работе использованы результаты, полученные в ходе выполнения проекта № 12-01-0130, выполненного в рамках Программы «Научный фонд НИУ ВШЭ» в 2013-2014 гг.

# Актуальность исследования



Ипотечный кризис

Несовершенство сложившейся  
практики управления  
кредитными рисками



Активная государственная  
жилищная политика и  
антикризисные меры по  
поддержке ипотечного  
кредитования

# Стилизованные факты

## 1. Агентство по ипотечному жилищному кредитованию

(АИЖК) :

- 100% акций принадлежит государству
- ипотечные кредиты, застрахованные государством
- система двухуровневого рефинансирования ИЖК
- разрабатывает кредитные программы
- рефинансирует риски своих региональных представительств

## 2. Высокорискованные заемщики.

3. Наличие квартальных лимитов у региональных представительств АИЖК.

#### 4. Вероятность дефолта (PD).

5. Дефолт – 90+ дней просроченной задолженности.

6. Отсутствие понятия «ипотечный дефолт» в российском законодательстве.

7. Ключевые драйверы дефолта:

- социально-демографические характеристики
- параметры ипотечного займа
- обслуживание ипотечного займа
- макроэкономическая ситуация

8. Модели бинарного выбора для оценки PD на рынке ипотечного кредитования страдают от проблемы самоотбора выборки (sample selection bias).

## 9. Оценки PD смещены и несостоятельны вследствие:

- проблемы самоотбора
  - смещение вследствие одновременности (не учитывается процесс кредитного андеррайтинга)
  - цензурированность (частичная наблюдаемость)
  - величина смещения зависит от степени корреляции процессов дефолта и кредитного андеррайтинга
- проблемы эндогенности объясняющих переменных
  - пропущенные существенные переменные
  - ошибки измерения регрессоров
  - проблема самоотбора
  - одновременные связи между переменными
  - серийная корреляция ошибок при наличии лагированной зависимой переменной среди регрессоров

# Проблема самоотбора

## Модель Хекмана (1976, 1979)

$$y_i = x_i' \beta + \varepsilon_i$$

Предположим, что мы наблюдаем  $y_i$  только при условии, когда ненаблюдаемая латентная переменная  $z_i^*$  превышает пороговое значение:

$$\text{Уравнение выхода} \quad \left\{ \begin{array}{l} y_i = x_i' \beta + \varepsilon_i, z_i^* > 0, \quad \varepsilon_i \sim N(0, \sigma^2) \\ y_i = \text{не наблюдаемо в противном случае.} \end{array} \right.$$

$$\text{Уравнение участия} \quad \left\{ \begin{array}{l} z_i^* = w_i' \alpha + u_i \\ z_i = \left\{ \begin{array}{l} 1, \text{ if } z_i^* > 0 \\ 0, \text{ в противном случае.} \end{array} \right. \end{array} \right. \quad \begin{array}{l} u_i \sim N(0, 1) \\ \text{corr}(\varepsilon_i, u_i) = \rho_{\varepsilon u} \end{array}$$

$$E(y_i | y_i \text{ наблюдаемо}) = E(y_i | z_i^* > 0) = \\ = x_i' \beta + E[\varepsilon_i | u_i > -w_i' \alpha] = x_i' \beta + \rho_{\varepsilon u} \sigma_\varepsilon \lambda_i(w_i' \alpha)$$

**$\lambda$  Хекмана (Обратное соотношение Миллса)**

$$\lambda_i(w_i' \alpha) = \frac{\varphi(w_i' \alpha)}{\Phi(w_i' \alpha)}$$

**Коэффициент перед  $\lambda$  указывает на наличие смещения вследствие самоотбора (sample selection bias)**

- 1) 2-х шаговая процедура Хекмана
- 2) Вариант оценивания с помощью MLE



# Проблема эндогенности

Attanasio, Goldberg, Kyriazidou (2008, International Economic Review), Das, Newey, Vella (2003, The Review of Economic Studies):

- **Оценка спроса на автокредиты с учетом проблемы эндогенности и самоотбора**
- **3-х шаговая процедура оценивания:**
  - оценивание уравнение участия
  - оценивание уравнений для эндогенных переменных с помощью полупараметрической регрессии с коррекция на проблему самоотбора
  - замена эндогенных переменных в уравнении спроса на их прогнозные значения и оценивание уравнение спроса с помощью полупараметрической регрессии

## Follain (1990, AREUEA Journal)

- **Ипотечный займ как последовательность решений заемщика и кредитной организации**
- Решения заемщика: выбор LTV, обслуживание согласно условиям кредитного договора, рефинансирование ссуда или объявление дефолта, включая выбор времени, а также выбор кредитного продукта
- **Самоотбор заемщиков**

## Rachils, Yezer (1993, Journal of Housing Research)

- 4 решения: выбор кредитной организации, выбор ипотечного продукта (LTV, срок), решение кредитора об одобрении/отклонении заявки, решение заемщика об обслуживании сделки (в т.ч. дефолт)
- **Решения должны рассматриваться как взаимозависимые**

Phillips, Yezer, Trost (1994, 1996, Journal of Real Estate Finance and Economics), Ross (2000, Journal of Real Estate Economics), Bajari, Chu, Park (2008, National Bureau of Economic Research) etc. :

- **Эмпирическое** показали, что моделирование процесса кредитного андеррайтинга и дефолта отдельно ведет к **смещенным оценкам параметров**

Phillips, Yezer (1996, Journal of Housing Research), Ross, Yinger (1999, Mortgage lending discrimination: A review of existing evidence), Ambrose et al. (2008, Journal of Real Estate Economics) etc.:

- **Эмпирическое подтверждение эндогенности** параметров кредитной сделки

Полтерович, Старков (2007, (Экономика и математические методы):

- Стратегии формирования ипотечного рынка в России
- Модельные расчеты по стратегии создания **массовой ипотеки** в России, основанной на трансплантации модифицированного **института стройсберкасс**

Ниворожкина, Овчарова, Синявская (2013, Прикладная эконометрика):

- Эконометрическое моделирование риска выплат по потребительским кредитам
- Опросные данные о кредитном поведении за 2011 г. проекта «Родители и дети, мужчины и женщины в семье и обществе»
- Статистически значимое влияние **социально-демографических характеристик и типа региона** на вероятность просрочки по выплате кредитов
- **Эмпирическое подтверждение**, что моделирование риска просрочки и вероятности иметь кредит отдельно ведет к **смещенным оценкам параметров ( $\rho=0,22^{***}$ )**

# Процесс ипотечного кредитования это последовательность решений:

1. Подача заявки заемщиком в кредитную организацию
2. Решение кредитной организации об одобрении/отклонении заявки
3. Установление величины кредитного лимита
4. Решение заемщика о заключении/не заключении кредитного контракта
5. Выбор параметров кредита и объекта недвижимости
6. Выбор стратегии обслуживания долга

# Исследовательские вопросы

1. Что является ключевыми драйверами процесса принятия решения агентов на рынке ипотечного кредитования? (займодаделец и заемщик)
2. Существует ли корреляция между разными шагами процесса ипотечного кредитования?
  - Обслуживание долга и подача заявки
  - Обслуживание долга и одобрение/отклонение заявки
  - Подача заявки и одобрение/отклонение заявки

- 1. Агрегированные региональные ежемесячные данные, по рынку ипотечных жилищных кредитов, застрахованных государством и региональные макроэкономические показатели за период 01.08.2008–31.08.2012.**
- 2. Данные одного из региональных представительств АИЖК по 4300 российским заемщикам, 2008–2012 гг.: социально-демографические характеристики, параметры ипотечного договора, характеристики предмета ипотеки, информация по обслуживанию ипотечного займа.**
  - 600 (13,95%) отклоненные заявки
  - 3700 (86,05%) одобренные заявки (выданные займы)
    - 2801 (74,55%) выданные ипотечные займы
      - 166 (5,93%) дефолты (90+)
      - 2635 (94,07%) не дефолты
    - 956 (25,45%) не выданные ипотечные займы

# Описательные статистики

Переменные	Описание	Mean	Std. Dev.	Min	Max
Age of borrower	Возраст клиента, лет	34	7,6	21	61
# of co-borrowers	Количество созаемщиков	0,62	0,57	0	3
Loan limit	Максимальная сумма кредитного лимита, руб.	936059,4	684952	0	12700000
Rate	Процентная ставка по кредиту, %	11,58	1,62	9,55	19
Loan amount	Сумма ипотечного займа, руб.	1039966	573503,1	120000	10000000
Down payment	Размер первоначального взноса, руб.	854494,6	706638,9	0	13800000
LTV	Соотношение суммы ипотечного займа к оценочной стоимости приобретаемого жилья	0,56	0,17	0,02	0,94
DTI	Соотношение размера ежемесячного платежа к ежемесячному доходу заемщика	0,45	0,18	0,06	1
Flat value	Оценочная стоимость приобретаемого жилья, руб.	1894460	1049331	330000	15300000
Days of observation	Наблюдаемое количество дней кредита, дней	786,65	430,77	15	1487
Unemployment rate	Квартальный уровень безработицы в регионе, %	8,43	1,51	6,3	10,9
Mean loan	Средний размер ипотечного займа в регионе, руб.	1160,27	252,23	899,31	1908,2



# Описательные статистики

Переменные	Описание	Mean	Std. Dev.	Min	Max
Median maturity	Медианный срок ипотечного займа в регионе, мес.	201,64	12,7	173	222,2
Median rate	Медианная процентная ставка по ипотечному кредиту в регионе, %	13,1	0,82	12	14,3
Mean DTI	Среднее соотношение размера ежемесячного платежа к ежемесячному доходу заемщика в регионе	34,81	0,7	33,44	36,68
Mean m2 value	Средняя стоимость 1 м <sup>2</sup> в регионе, руб.	37617,03	6395,77	28782	51304
Lodging coefficient in years	Коэффициент доступности жилья, лет	3,48	0,68	2,57	4,65
Mortgage volume	Объем выданных ипотечных кредитов в регионе, млн. руб.	885948,4	563161,6	116100	2191000
Mortgage amount	Количество выданных ипотечных кредитов в регионе, шт.	896,57	528,89	134	2112

После исключения выбросов (597 наблюдений):

- 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, равным 0 или превышающим.

<b>Переменные</b>	<b>Всего</b>	<b>Доля, %</b>
<b>Пол</b>		
мужской	1881	43,74
женский	2419	56,26
	4300	100
<b>Семейное положение</b>		
не замужем/холостяк	1221	28,70
замужем/женат	2359	55,45
вдовец/вдова	56	1,32
в разводе	618	14,53
	4254	100
<b>Место работы</b>		
не работает	1	0,02
пенсионер	0	0,00
военнослужащий	13	0,31
наемный работник	3965	95,27
индивидуальный предприниматель	39	0,94
госслужащий	144	3,46
	4162	100
<b>Уровень образования</b>		
начальное	65	1,59
среднее	1748	42,69
незаконченное высшее	138	3,37
высшее	2144	52,36
	4095	100

<b>Переменные</b>	<b>Всего</b>	<b>Доля, %</b>
Ежемесячный доход заемщика		
данные о доходе отсутствуют	2918	67,86
0-9999 руб.	118	2,74
10000-19999 руб.	376	8,74
20000-39999 руб.	599	13,93
>=40000 руб.	289	6,72
	4300	100
Сумма основного ежемесячного дохода созаемщиков		
данные о доходе отсутствуют	3725	86,63
0-9999 руб.	160	3,72
10000-19999 руб.	225	5,23
>=20000 руб.	190	4,42
	4300	100
Тип процентной ставки		
фиксированная	2423	86,50
плавающая	378	13,50
	2801	100
Срок кредита		
< 120 мес.	181	6,46
120-179 мес.	595	21,25
180-239 мес.	1107	39,54
240-299 мес.	690	24,64
>=300 мес.	227	8,11
	2800	100

## 1. Моделирование вероятности подачи заявки на получение ИЖК на агрегированных данных

$$y_{1it} = g_{1t}(z_{1t}) + e_{1it}$$
$$e_{1it} \sim IID(0, \sigma_{e_{it}}^2)$$
$$z_{1t} \cong (D_t, M_t),$$

$t = 1, \dots, T, T$  – количество месяцев

$y_{1it}$  – вероятность подачи заявки  $i$ -м потенциальным заемщиком в момент времени  $t$  (количество поданных заявок в месяц, деленное на число домашних хозяйств, одинаковая для заемщиков. Подающих заявку в один месяц)

$D_t$  – вектор агрегированных демографических характеристик

$M_t$  – вектор экзогенных макроэкономических показателей

## 2. Моделирование вероятности одобрения кредитной заявки для всех поданных заявок

$$y_{2it}^* = g_2(z_{*2it}^*) + e_{2i}$$

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

$$\text{cov}(e_{1it}, e_{2i}) = \sigma_{12}, i \in N_t$$

$$\text{cov}(e_{1it}, e_{2i}) = 0, i \notin N_t$$

$$z_{2it}^* = (D_i^*, M_t)$$

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

## 3. Моделирование величины кредитного лимита

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

$$\bar{L}_{it} = y_{2it} \bar{L}_{it}^* - \text{наблюдаемо}$$

$$\text{cov}(e_{2i}, e_{\bar{L}i}) = \sigma_{2\bar{L}}$$

$\bar{L}$  – величина кредитного лимита

$z_{\bar{L}it}^*$  – вектор инструментальных демографических характеристик и макроэкономических показателей на дату подачи кредитной заявки

## 4. Моделирование вероятности заключения кредитного договора

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

$$y_{3it} = y_{2it} y_{3it}^* - \text{наблюдаемо}$$

$$\text{cov}(e_{2i}, e_{3i}) = \sigma_{23}$$

$$z_{3it}^* = (D_i^*, M_t)$$

$$\hat{L}_{it}^* = \hat{g}_{\bar{L}}(z_{\bar{L}it}^*)$$

$y_{3i} = 1$  – 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

## 5. Одновременный выбор параметров ипотечного займа и стоимости приобретаемого жилья

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

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

$$\text{cov}(e_{3i}, e_{ji}) = \sigma_{3j}$$

$$\text{cov}(e_{3i}, e_{Vi}) = \sigma_{3V}$$

$$\text{cov}(e_{ji}, e_{ki}) = \sigma_{jk}$$

$$\text{cov}(e_{ji}, e_{Vi}) = \sigma_{jV}$$

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

$V$  – 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, макроэкономических показателей на дату подачи кредитной заявки, и характеристики приобретаемого жилья.



## 6. Моделирование вероятности дефолта

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

$$y_{4it} = y_{3it} y_{4it}^* - \text{наблюдаемо}$$

$$\text{cov}(e_{3i}, e_{4i}) = \sigma_{34}$$

$y_{4it}^*$  – вероятность дефолта

$\hat{C}_{it}^*$  – прогнозные значения параметров кредитного контракта

$\hat{V}_{it}^*$  – прогнозная стоимость приобретаемого жилья

$z_{4it}^*$  – вектор экзогенных индивидуальных демографических характеристик и макроэкономических показателей на дату подачи кредитной заявки

## Предположения:

1. Ошибки в уравнениях (1-6) имеют совместное распределение с нулевым вектором первых моментов.
2. Для корректировки на проблему самоотбора используется лямбда Хекмана (обратное соотношение Миллса)
3. Линейный  $g$ -функции используются для непрерывных зависимых переменных, пробит  $g$ -функции - для дискретных зависимых переменных.

# Эмпирические результаты

## 1. Вероятность подачи заявки на получение ИЖК

- Объясняется макроэкономическими показателями региона, характеризующими спрос и предложение на рынке ИЖК.
- Увеличивается в случае негативных шоков спроса (рост цен на недвижимость и снижение доходов населения) или предложения на рынке ипотеки (увеличение средних ставок, снижение средних сроков выдаваемых кредитов).
- В меньшей степени на спрос оказывают влияние параметры ипотечного займа и демографические характеристики клиентов.

## 2. Вероятность одобрения кредитной заявки

- Процесс кредитного андеррайтинга региональных отделений определяется не только характеристиками потенциальных заемщиков, но и количеством поданных кредитных заявок в конкретный период времени, которые отрицательно влияют на вероятность одобрения кредитной заявки.

## 3. Установление величины кредитного лимита

## 4. Вероятность заключения кредитного договора

- Статистически значимое положительное влияние оказывает семейное положение и место работы, а также сумма займа, величина выставленного кредитного лимита, срок ипотеки наряду с низкой процентной ставкой.

## 5. Выбор параметров ипотечного займа и стоимости приобретаемого жилья

- Спрос на ИЖК, застрахованные государством, представляет собой функцию от размера кредитного лимита, характеристик потенциального желаемого предмета ипотеки.
- В меньшей степени на спрос оказывают влияние параметры ипотечного займа и демографические характеристики клиентов.

## 6. Вероятность дефолта

### Примечание.

Базисные категории — заемщик наемный работник, ежемесячный доход 0-9999 руб., срок кредита >=300 месяцев.

\*\*\*, \*\*, \* — значимость на уровне 1, 5 и 10% соответственно.

	Коэффициенты
Вероятность подачи ипотечной заявки	0.061***
Вероятность одобрения ипотечной заявки	1.580*
Вероятность заключения ипотечного договора	0.510
Плавающая процентная ставка	-11.641
Размер процентной ставки	1.659***
Срок кредита <120 месяцев	8.310
Срок кредита 120-179 месяцев	15.2978**
Срок кредита 180-239 месяцев	14.128*
Срок кредита 240-299 месяцев	29.278***
Оценочная стоимость жилья, тыс. руб.	-0.009***
Сумма займа, тыс. руб.	0.018***
LTV	-36.063**
Возраст заемщика	-0.742**
Квадрат возраста заемщика	0.0106*
Мужской пол	0.255**
Заемщик государственный служащий	0.776
Количество созаемщиков	-0.349***
Отсутствуют данные о ежемесячном доходе	0.392
Ежемесячный доход 10000-19999 руб.	-0.427*
Ежемесячный доход 20000-39999 руб.	-0.080
Ежемесячный доход >=40000 руб.	-0.130
Средняя стоимость 1 м <sup>2</sup> в регионе, тыс. руб.	-0.054**
Наблюдаемое количество дней кредита	0.0005
Константа	-9.452

# Эмпирические результаты

- PD выше, чем выше размер процентной ставки, сумма ипотечного займа, срок кредита.
- PD выше для заемщиков со средним уровнем дохода, среднего возраста и с меньшим числом созаемщиков.
- Параболической зависимости PD и возраста заемщика.
- Статистически значимая положительная корреляции между ошибками в уравнениях PD и подачи заявки в АИЖК: в моменты негативных шоков ипотечного рынка, кредитная организация выбирает стратегию доступных цен на ипотеку (при росте рыночных ставок и снижении общих объемов рынка увеличивается вероятность подачи в АИЖК), выполняя социальную программу, вопреки росту рыночного риска.
- Статистически значимая положительная корреляции между ошибками в уравнениях PD и одобрения заемщика: одобрение высокорискованных заемщиков, выполнение социальных функций АИЖК по обеспечению доступным жильем заемщиков

# Выводы

1. Предложенная многошаговая процедура оценивания позволяет получить состоятельные оценки PD в результате корректировки на эндогенность и самоотбор.
2. Существует корреляция между разными этапами процесса ипотечного кредитования – необходимо их совместное моделирование, что позволяет повысить точность прогноза.
3. Заемщики по ипотечным жилищным кредитам, выдаваемым АИЖК, характеризуются более высоким уровнем кредитного риска (? *Страхование рисков*)
4. Полученные результаты могут служить основой для «настройки» действующих систем риск-менеджмента в кредитной организации.



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

Спасибо  
за внимание!

[AMPoroshina@gmail.com](mailto:AMPoroshina@gmail.com)

ул. Лебедева 27, Пермь, Россия, 614000

[www.hse.ru](http://www.hse.ru)