

ЭКОНОМИЧЕСКОЕ ОБОЗРЕНИЕ
Национального Банка Республики Казахстан

№ 2, 2021

Нур-Султан, 2021 год

ЭКОНОМИЧЕСКОЕ ОБОЗРЕНИЕ Национального Банка Республики Казахстан

Издатель: Национальный Банк Республики Казахстан

Редакционная коллегия издания

Председатель редакционной коллегии

Заместитель Председателя Национального Банка Баймагамбетов А.М.

Члены редколлегии:

Руководитель подразделения денежно-кредитной политики

Руководитель подразделения финансовой стабильности и исследований

Руководитель подразделения платежного баланса

Руководитель подразделения монетарных операций

Руководитель подразделения развития финансовых организаций

Ответственный за выпуск издания – работник Центра исследований и аналитики

**Точка зрения и мнения авторов статей не являются официальной позицией
Национального Банка Республики Казахстан и могут не совпадать с ней.**

ISSN 2709-2496

ЭКОНОМИЧЕСКОЕ ОБОЗРЕНИЕ
Национального Банка Республики Казахстан

№ 2, 2021

СОДЕРЖАНИЕ

Анализ рисков потребительских кредитов с помощью алгоритмов машинного обучения <i>Байкулаков Ш.Б., Белгибаев З.Э.</i>	4
Ликвидность рынка ГЦБ: проблемы и перспективы решений <i>Алимбетова И.С., Кадырбеков Т.Ж., Мустафин Е.Т., Ыбраев Ж.Ж.</i>	27
Перспективы и развитие небанковских финансовых организаций <i>Бакиров А.М., Бижанов Б.А., Завалина А.В., Садыбаева Д.С.</i>	46

АНАЛИЗ РИСКОВ ПОТРЕБИТЕЛЬСКИХ КРЕДИТОВ С ПОМОЩЬЮ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Байкулаков Ш.Б. – директор проектного офиса, Центр развития платежных и финансовых технологий Национального Банка Республики Казахстан.

Белгибаев З.Э. – главный специалист-аналитик управления монетарного анализа департамента денежно-кредитной политики Национального Банка Республики Казахстан.

Данное исследование представляет собой попытку оценки кредитоспособности физических лиц с помощью алгоритмов машинного обучения на основе данных, предоставляемых банками второго уровня Национальному Банку Республики Казахстан. Оценка кредитоспособности заемщиков позволяет НБРК исследовать качество выданных кредитов банками второго уровня и прогнозировать потенциальные системные риски.

В данном исследовании были применены два линейных и шесть нелинейных методов классификации (линейные модели - логистическая регрессия, стохастический градиентный спуск, и нелинейные - нейронные сети, k-ближайшие соседи (kNN), дерево решений (decision tree), случайный лес (random tree), XGBoost, наивный Байесовский классификатор (Naïve Bayes)) и сравнивались алгоритмы, основанные на правильности классификации (accuracy), точности (precision) и ряде других показателей. Нелинейные модели показывают более точные прогнозы по сравнению с линейными моделями. В частности, нелинейные модели, такие как случайный лес (random forest) и k-ближайшие соседи (kNN) на передискредитированных данных (oversampled data) продемонстрировали наиболее многообещающие результаты.

Ключевые слова: потребительские кредиты, машинное обучение, банковское регулирование, стохастический градиентный спуск, логистическая регрессия, k-ближайшие соседи, классификатор случайных лесов, дерево решений, gaussian NB (Гауссовский наивный Байесовский классификатор), XGBoost, нейронные сети (многослойный перцептрон).

JEL-классификация: G21, G28, E37, E51.

1. Введение

Расширение потребительских кредитов является основным фактором роста розничных кредитов в последние годы. Выдача потребительских кредитов может привести к экономическому росту в стране, но в то же время следует отметить, что чрезмерное финансирование банками может вызвать появление системных рисков. Быстрый рост потребительского кредитования можно объяснить появлением детальных данных о физических лицах и эффективных инструментов оценки кредитного риска с одной стороны, а также стабилизацией экономической ситуации и ростом благосостояния населения – с другой.

Динамичное расширение потребительского кредитования несет в себе ряд системных рисков, связанных, прежде всего, с ростом уровня долговой нагрузки на отдельные слои населения. В случае падения реальных доходов населения банковская система может столкнуться с массовыми дефолтами по потребительским кредитам. Как следствие, это может привести к снижению совокупного спроса в экономике, что в итоге может негативно повлиять на состояние экономики.

На основе изучения исследовательских работ по применению алгоритмов машинного обучения можно сделать вывод, что наиболее популярным методом анализа кредитных рисков потребительских кредитов на больших массивах данных является

использование нелинейных моделей, таких как нейронные сети, k-ближайшие соседи, дерево решений, случайный лес, XGBoost, наивный байесовский классификатор, а также такие линейные модели, как логистическая регрессия, стохастический градиентный спуск. Большинство авторов использовали данные кредитных бюро и банков второго уровня. Учитывая, что данное исследование проводилось на основе регуляторных данных, собранных центральным банком, могут быть некоторые расхождения в подходах с возможными минимальными отличиями в результатах.

Первый раздел представляет собой обзор литературы, в котором рассматриваются аналогичные работы других авторов. Во втором разделе описывается методология исследования, а также список использованных метрик. Далее следует раздел обсуждения результатов, в котором авторы описывают результаты моделирования. В заключительном разделе работы описываются основные выводы исследования.

2. Обзор литературы

По результатам исследования Grier (2012) было выявлено, что для анализа потребительского кредитования необходимо учитывать 5 параметров [13]. Первый параметр – это репутация заемщика. В настоящее время кредитные менеджеры имеют доступ к отчетам кредитного бюро, которые показывают кредитную историю потребителя до принятия решения о выдаче кредита. Второй фактор – это платежеспособность потребителя, связанная с его доходами и существующими финансовыми обязательствами. Капитал – третий параметр, указывающий на первоначальный взнос, который может себе позволить заемщик. Под четвертым фактором подразумевают внешние факторы, такие как состояние рынка труда и общие экономические условия. Последним параметром является залог.

Методы кредитного скоринга оценивают вероятность погашения кредита потребителем, используя информацию из отчета кредитного бюро и кредитной заявки [13]. Скоринг приложений – это первый тип модели кредитного скоринга, который оценивает заявки новых потребителей на основе таких параметров, как капитал, мощность и т. д. Другая модель, такая как модель оценки поведения, оценивает платежеспособность потребителей, предупреждая о потенциальных просроченных счетах.

Одна из целей использования метода кредитного скоринга – снизить риски невыплат по выдаваемым кредитам в будущем. Помимо традиционных методов, банки недавно начали интегрировать алгоритмы машинного обучения (ML) в кредитный скоринг. Так, Henley and Hand (1996) сравнили методы машинного обучения, такие как метод классификации k-ближайшего соседа (kNN), с традиционными методами кредитного скоринга, такими как линейная, логистическая регрессия и дерево решений [15]. Алгоритмы были протестированы на предмет риска безнадёжных долгов, и лучшим методом был определен метод с наименьшим ожидаемым уровнем риска. Несмотря на нечувствительность классификации k-ближайшего соседа к параметрам, она дала лучший результат. Кроме того, для проведения оценки потребителя на предмет платежеспособности по кредиту требуется несколько секунд [15].

Далее, Henley, Gueran, and Hassani (2018) применили логистическую регрессию, классификатор случайного леса и повышения градиента, а также модель глубокого обучения для анализа кредитного риска компаний. Были использованы различные наборы данных, а затем было выбрано 10 наиболее важных параметров и те же методы использовались для сравнения результатов. Предполагается, что лучший алгоритм показывает наибольшую площадь под кривой (AUC) и наименьшую среднеквадратичную ошибку (RMSE). Бинарные классификаторы превзошли модели глубокого обучения, и лучшая производительность принадлежит классификатору, такому как градиентный бустинг.

Для оценки заявок на получение потребительского кредита также используются регрессионные модели машинного обучения. Munkhdalai et al. (2019) сравнили алгоритмы

машинного обучения с такими моделями, как система кредитного рейтинга FICO [20]. Данные Survey of Consumer Finances (SCF) были использованы в качестве набора данных, к которым были применены различные методы регрессии машинного обучения. При этом глубокие нейронные сети и алгоритмы XGBoost показали более высокую точность. В результате определили, что кредитные убытки были бы ниже, если бы кредитные учреждения начали использовать модели машинного обучения с 2001 года. Кроме того, глубокие нейронные сети и алгоритм xgboost показали более высокую точность.

Brown and Mues (2012) протестировали пригодность и точность методов классификации, таких как логистическая регрессия, нейронная сеть, дерево решений, градиентный бустинг, метод опорных векторов с квадратичной функцией потерь (LS-SVM) и случайный лес, для несбалансированного набора данных о ссуде [4]. К набору данных был применен метод недостаточной выборки, а затем дисбаланс набора данных постепенно увеличивался для оценки методов классификации машинного обучения. Результат показывает, что классификаторы случайного леса и градиента бустинга хорошо справляются с несбалансированными данными, в то время как дерево решений, квадратичный дискриминантный анализ (QDA) и k-ближайшие соседи (kNN) работают хуже, чем другие методы.

Помимо вышеперечисленных моделей, в машинном обучении также есть несколько алгоритмов. Baesens et al. (2003) протестировали метод опорных векторов на основе ядра (kernel-based SVM) и метод опорных векторов с квадратичной функцией потерь (LS-SVM), а также другие популярные классификаторы машинного обучения для реальных наборов данных кредитного скоринга, включая наборы данных финансовых учреждений Бенилюкса и Великобритании [3]. Результат показывает, что классификатор нейронной сети и машины векторов поддержки на основе ядра работали очень хорошо. Автор также отметил хорошую производительность линейного дискриминантного анализа и логистической регрессии. 41 классификатор, включая новые методы оценки кредитоспособности, применяемые к одним и тем же наборам данных (Lesmann, Baesens, Seow, and Thomas, 2015). Результат исследования показывает, что искусственные нейронные сети (ИНС) работают лучше, чем экстремальные методы обучения (ELM), и случайный лес (RF) лучше, чем ротационный лес (RotFor). Однако эти методы не могут дать экономическую интерпретацию модели, и для достижения этой цели необходимо провести дальнейшие исследования. Классификатор случайного леса был выбран в качестве эталона, так как он может дать пояснительную информацию для фундаментального анализа.

Tsai and Chen (2010) разработали четыре гибридных метода машинного обучения для сравнения с простыми классификаторами [22]. Гибридный алгоритм – это комбинация двух методов машинного обучения. В этом исследовании были выбраны методы классификации и кластеризации, а также четыре различных метода. Результат показывает, что комбинация логистической регрессии (LR) и нейронных сетей показала самый высокий прогноз, в то время как «двойная кластеризация» была худшим алгоритмом.

В 2015 году участники реализовали XGBoost в 17 из 29 решений Kaggle. Алгоритм был реализован для прогнозирования продаж магазина, классификации веб-текста, прогнозирования поведения клиентов, классификации вредоносных программ [9]. В этом исследовании алгоритм будет использоваться для кредитного анализа и сравниваться с другими методами.

В этом исследовании к набору данных были применены два линейных и шесть нелинейных методов классификации, которые будут описаны в следующем разделе. Алгоритмы сравнивались по правильности классификации (accuracy), точности (precision) и нескольким другим показателям.

3. Методология исследования и исходные данные

Для анализа портфелей потребительских кредитов казахстанских банков использовались алгоритмы машинного обучения. Данные получены из Кредитного регистра, формируемого на основе информации, предоставленной банками второго уровня Национальному Банку Казахстана (НБК). Данные очищены от кредитов с недостающими и ненадежными параметрами. Кроме того, из данных были удалены ссуды с просрочкой платежа не более 90 дней. Если просрочка платежа по кредиту превышает 90 дней, ссуда признается неработающей ссудой (NPL). Работающие кредиты не имеют просрочки платежа со стороны потребителя.

Таблица 1

Параметры данных для исследования

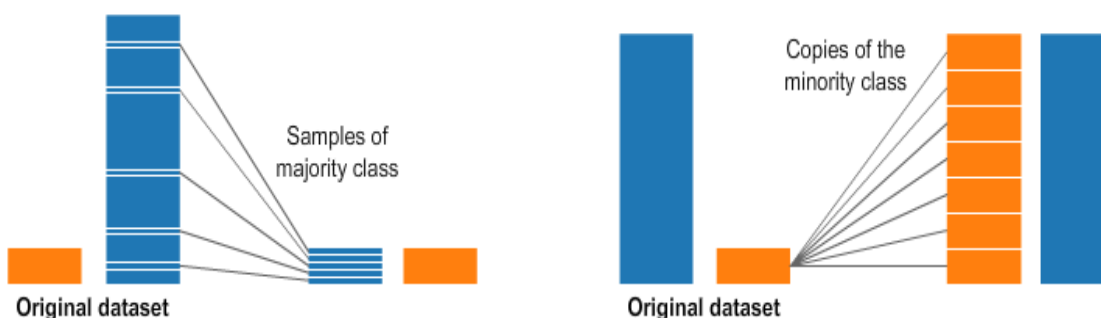
Параметры	Виды
Регионы	Нур-Султан, Алматы и 15 регионов
Валюта	Тенге, рубль, долл. США и евро
Тип займа	Наличные займы и кредитная карта
Объект кредитования	Потребительские и автокредиты
Пол	Мужчина и женщина
Гражданство	Резидент и нерезидент
Процентная ставка	От 0% до 56%
Сумма займа	От 10 000 до 15 млн тенге
Возраст	От 18 до 99 лет

Источник: Национальный Банк Республики Казахстан

Стратегия семплинга для несбалансированных данных. Для повышения точности прикладных алгоритмов машинного обучения были использованы такие стратегии семплинга, как удаление примеров мажоритарного класса (субдискретизация (undersampling)) и увеличение числа примеров миноритарного класса (передискретизация (oversampling)). Поскольку количество хороших ссуд было значительно больше, чем проблемных ссуд, данные были сбалансированы и было проведено сравнение двух методов.

Рисунок 1

Методы ресемплинга для несбалансированных данных
Undersampling Oversampling



Источник: Alencar, 2017

Случайное удаление примеров мажоритарного класса (случайная субдискретизация (random undersampling)) и случайное дублирование примеров миноритарного класса (случайная передискретизация (random oversampling)) – самые популярные и простые стратегии семплинга. Первая стратегия создает новый класс с равным размером класса меньшинства, взятого из класса большинства, тогда как вторая стратегия дублирует класс

меньшинства несколько раз, пока длина классов большинства и меньшинства не станет равной. Недостатком случайного удаления примеров мажоритарного класса (random undersampling) является потеря информации. Однако это единственная подходящая стратегия недостаточной выборки для набора данных, состоящего из категориальных и числовых данных. Так как случайное дублирование примеров миноритарного класса (random oversampling) приводит к переобучению ввиду того, что он копирует класс меньшинства [2], был выбран другой, подходящий метод семплинга.

SMOTE (метод синтетической передискретизации меньшинства) – еще один популярный метод семплинга, который рисует новые выборки классов меньшинств, используя существующие данные. Он проводит линии между примером и ближайшими 5 соседями и создает новый искусственный образец вдоль этой линии. Рисунок 2 иллюстрирует, как создавались новые образцы, тем самым предоставляя дополнительную информацию для модели машинного обучения [7].

SMOTE нельзя применить к вышеупомянутому набору данных, поскольку он не применим к категориальным переменным. Поэтому в качестве стратегии семплинга выбран метод синтетического SMOTE-NC. Он работает как метод SMOTE для непрерывного набора данных, а категориальной переменной новой выборки является значение большинства k-ближайших соседей [8].

Рисунок 2



Источник: Lemaitre, Nogueira, Oliveira and Aridas, n.d.

Предварительная обработка (Preprocessing). Предварительная обработка – важная задача, позволяющая сделать данные применимыми. Шесть параметров набора данных являются категориальными данными, и их необходимо преобразовать в числовые с помощью кодировщика. Библиотека Scikit-learn (Python) предоставляет методы, которые преобразуют дискретные данные в простой числовой массив. В нем также есть модели, которые разделяют данные на наборы данных для обучения и тестирования [10].

Следующей важной частью предварительной обработки является шкалирование данных. Алгоритмы машинного обучения обычно используют евклидову меру расстояния. Таким образом они становятся чувствительными к величине параметров. Например, алгоритмы игнорируют другие параметры в наборе данных из-за больших значений суммы кредита. Следовательно, для нормализации набора данных требуется масштабирование функций. Это также снижает стоимость: методы кредитного скоринга быстрее анализируют нормализованный набор данных. Вышеупомянутая библиотека Python предоставляет стандартные методы шкалирования, которые нормализуют набор данных с помощью формулы:

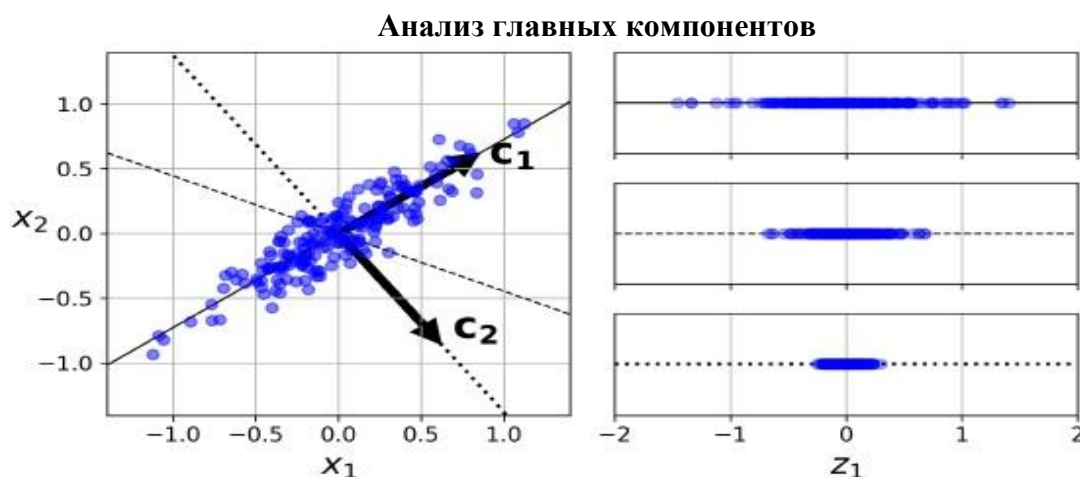
$$z = \frac{x - \mu}{\sigma}$$

где μ – среднее значение и σ является стандартной девиации данных [10].

Метод главных компонент (Principal component analysis). Метод главных компонент (РСА) – один из хорошо известных методов уменьшения размерности. Это помогает уменьшить размер набора данных с минимальной потерей информации [19]. Следовательно, это уменьшит время вычисления алгоритма.

Алгоритм выбирает правильную гиперплоскость и проецирует данные на эту гиперплоскость. После прогноза дисперсия данных вычисляется на основе каждой новой оси, называемой главным компонентом. Первый компонент сохраняет максимальную дисперсию, а последний компонент – наименьшую дисперсию.

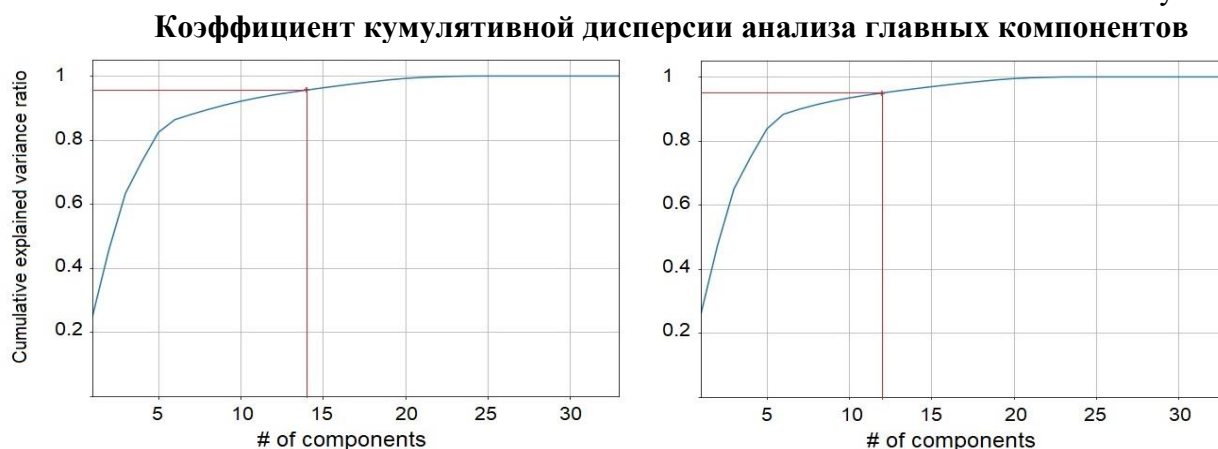
Рисунок 3



Источник: Герон, 2017

Метод главных компонент выполняется для обоих сбалансированных наборов данных. Результат субдискретизированных данных показывает, что первых 14 компонентов (столбцов) достаточно, чтобы охватывать 95% общей дисперсии набора данных, а остальные 19 компонентов можно удалить. Достаточно использовать только 12 компонентов набора передискретизированных данных, чтобы охватывать 95% общей дисперсии.

Рисунок 4



Источник: составлено авторами

В этом исследовании к набору данных были применены 2 линейных и 6 нелинейных методов машинного обучения. Основная цель – выяснить, какой алгоритм

превосходит другие. Кроме того, данное исследование помогает сравнивать и противопоставлять линейные и нелинейные алгоритмы.

Возможность обработки большого набора данных – главный критерий выбора алгоритма. Поэтому известные алгоритмы, такие как линейный классификатор и классификатор опорных векторов (SVC), в исследовании не рассматривались. Для целей анализа были использованы логистическая регрессия, классификатор стохастического градиентного спуска (SGD), наивный байесовский классификатор, k-ближайшие соседи (kNN), дерево решений (decision tree), случайный лес (random tree), классификатор многослойного персептрона (MLP), классификатор нейронной сети и XGBoost и обсуждались результаты. Некоторые гиперпараметры нелинейных алгоритмов не будут учитываться из-за невозможности обработки больших наборов данных.

Логистическая регрессия (Logistic Regression). Несмотря на название, логистическая регрессия используется для классификации. Алгоритм рассчитывает вероятность на основе обучающего набора данных по данной формуле:

$$P(y = 1|x) = \frac{1}{1+e^{-(w_0+W^T x)}},$$

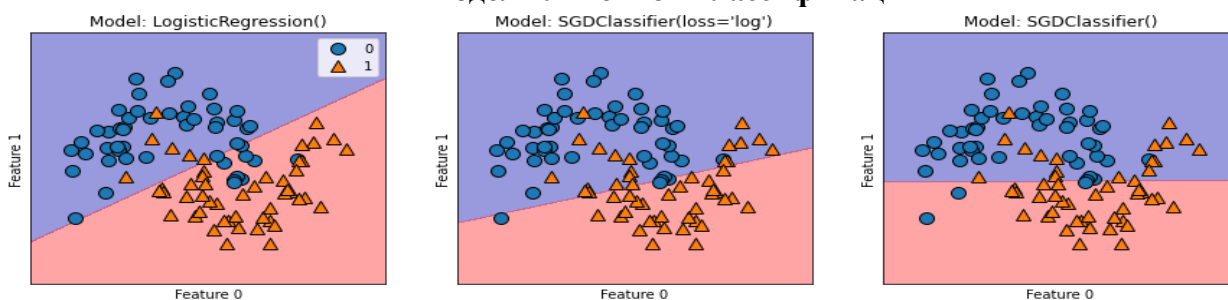
где x – входные данные, w_0 – скалярный вектор пересечения, а W – вектор параметров [3]. Если вероятность больше 50%, входные данные классифицируются как положительные.

Для повышения точности были настроены значимые гиперпараметры, такие как параметры решателя (solver) и регуляризации. Решатель Liblinear обычно используется для небольшого набора данных, тогда как sag и saga – лучший выбор для анализа более крупных данных, поскольку он требует меньше времени для вычислений [10].

Штраф (penalty) – это гиперпараметр регуляризации, используемый при наложении штрафа, и он имеет тесную связь с решателем. Решатели Newton-cg, sag и lbfgs поддерживают только штрафы «l2», в то время как решатель saga поддерживает штраф elasticnet. C – это величина, обратная силе регуляризации, которая должна быть положительной. Меньшее значение C указывает на более сильную регуляризацию [10].

Рисунок 5

Модели линейной классификации



Источник: составлено авторами

Стохастический градиентный спуск (Stochastic Gradient Descent (SGD)). Классификатор SGD – один из эффективных методов линейной классификации, применяемых к большим наборам данных. Алгоритм использует процедуру обучения SGD первого порядка. Параметр метода обновляется итеративно на обучающих примерах по данной формуле:

$$\omega = \omega - \eta \left[\alpha \frac{\partial R(\omega)}{\partial \omega} + \frac{\partial L(\omega^T x_i + b, y_i)}{\partial \omega} \right]$$

где α – гиперпараметр, который контролирует силу регуляризации, R – выражение регуляризации, который снижает сложность модели, L – функция потерь, которая измеряет соответствие модели, η – скорость обучения, а b – точка пересечения, которая обновляется аналогичным образом без регуляризации [10].

Градиентный спуск – один из важных алгоритмов, используемых для минимизации функции стоимости [11]. В целом существует три популярных типа градиентного спуска:

1. Общий градиентный спуск (Batch gradient descent);
2. Стохастический градиентный спуск (Stochastic gradient descent);
3. Градиентный спуск mini-batch (Mini-batch gradient descent).

Общий градиентный спуск вычисляет частную производную функции стоимости алгоритма с учетом его параметров. Другими словами, алгоритм обнаруживает, как различные значения параметров модели влияют на функцию стоимости алгоритма. Недостатком является то, что он использует целые обучающие данные для их вычисления на каждом этапе. Следовательно, алгоритм не может обрабатывать большие наборы данных [12].

Стохастический градиентный спуск решает эту проблему, поскольку он выбирает случайные экземпляры обучающих данных и вычисляет градиентный спуск на основе этого единственного экземпляра. С другой стороны, пакетный градиентный спуск будет плавно уменьшать функцию стоимости, в то время как она будет колебаться вверх и вниз, пока алгоритм не остановится. Алгоритм не может дать оптимальные значения параметров [12].

Градиентный спуск mini-batch берет небольшую выборку из обучающих данных и вычисляет алгоритм общего градиентного спуска. Следовательно, результат вычисления функции стоимости менее ошибочен по сравнению со стохастическим градиентным спуском [12].

Метод имеет много гиперпараметров, что делает его очень сложным. Важные параметры, такие как функция потерь (loss function) и параметры регуляризации, такие как штраф (penalty) и альфа (alpha), учитываются в процессе настройки модели. Метод с функцией потерь Hinge работает как линейная SVM, а с функцией log оперирует как логистическая регрессия [10].

Алгоритм очень чувствителен к шкалированию. Однако его можно легко реализовать, и он очень эффективен, особенно для больших наборов данных [10]. Кроме того, он продолжает работать без сохранения записи в ОЗУ [11].

Наивный байесовский классификатор (Naïve Bayes classifier). В общем классификатор наивного байеса (NB) основан на теореме Байеса, предполагающей независимость каждого параметра:

$$P(y|x_1, \dots, x_n) = \frac{P(y) \prod_{i=1}^n P(x_i|y)}{P(x_1, \dots, x_n)} [10]$$

Есть несколько типов «наивных байесовских классификаторов». В этом исследовании Гауссовский наивный Байесовский классификатор (Gaussian Naïve Bayes (GaussianNB)) используется для оценки вероятности признаков на основе данной формулы:

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(x_i-\mu_y)^2}{2\sigma_y^2}}$$

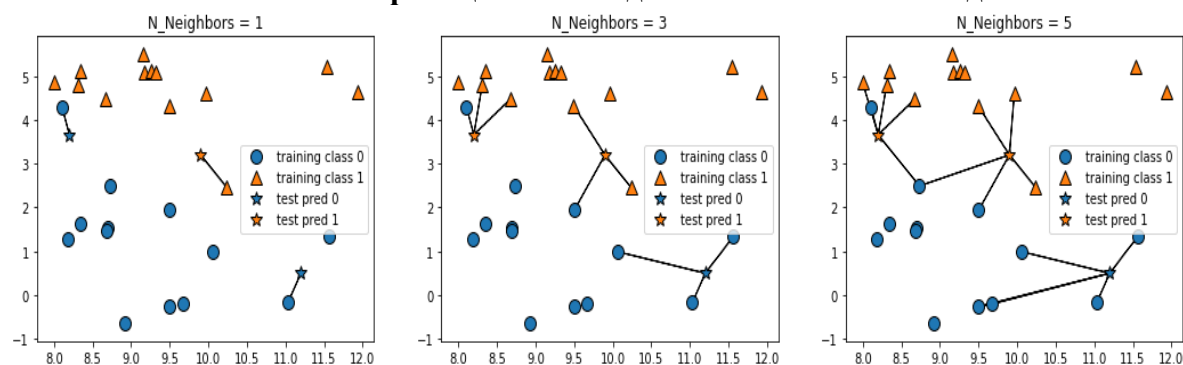
где среднее значение и стандартное отклонение оцениваются через максимальное правдоподобие [10].

k-ближайшие соседи (kth Nearest Neighbors (kNN)). k-ближайшие соседи – один из простейших алгоритмов машинного обучения. Количество ближайших соседей (k) является основным параметром в этом алгоритме. Результат новых данных будет идентифицирован на основе большинства результатов ближайших соседей. Большое значение k подавит влияние шума, но будет на границе классификации [10]. Расстояние измеряется по формуле евклидова расстояния:

$$d(x_i, x_j) = \|x_i - x_j\| = \left[(x_i - x_j)^T (x_i - x_j) \right]^{1/2} [4]$$

Рисунок 6

Классификационная модель k-ближайших соседей



Источник: Mlglearn library (Mueller & Guido, 2017)

Дерево решений (Decision Tree classifier). Дерево решений является простым алгоритмом, который прогнозирует целевой результат путем применения правил принятия решений на основе данных функций. Сильной стороной метода является способность обрабатывать числовые, категориальные характеристики и задачи с несколькими выходами. Еще одно преимущество – простота модели. Но модель может создавать чрезмерно сложные деревья, которые приводят к переобучению модели [10].

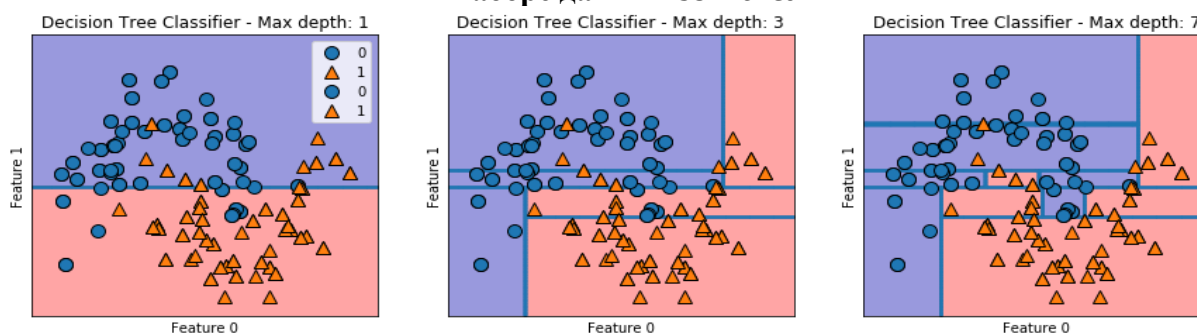
Критерий и максимальная глубина дерева учитываются в процессе настройки модели. Первый параметр – это функция, которая измеряет качество разделения. Gini (Gini impurity) и entropy (information gain) – два варианта выбора для функции критериев [10].

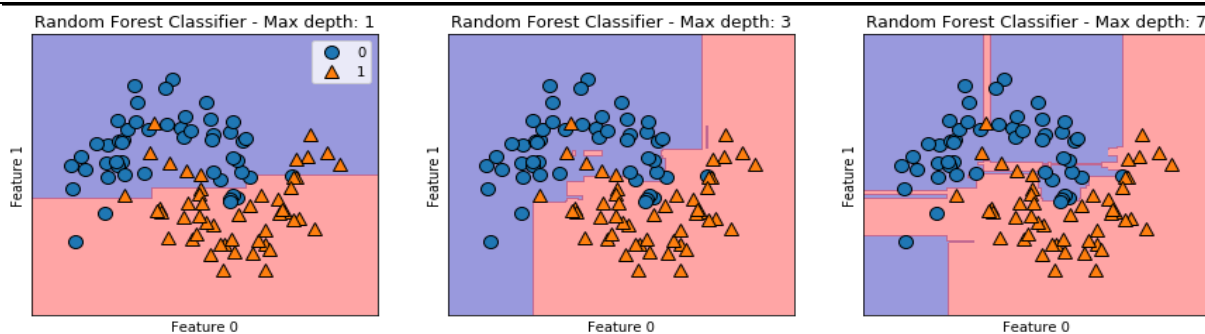
Случайный лес (Random Forest Classifier). Классификатор случайного леса – это метод классификации ансамбля, который имеет все гиперпараметры классификатора дерева решений. Разница между двумя моделями заключается в том, что первая ищет лучшую функцию среди подмножества функций, а вторая ищет лучшую функцию при разбиении узла. Таким образом, случайный классификатор леса приводит к большему дереву [12].

Классификатор дерева решений показывает более высокую дисперсию, что приводит к переобучению. Классификатор случайных лесов объединяет разнообразные деревья, тем самым добиваясь уменьшения дисперсии. Впоследствии он создает лучшую модель, чем классификатор дерева решений [10].

Рисунок 7

Дерево решений и классификатор случайного леса в наборе данных scikit-learn



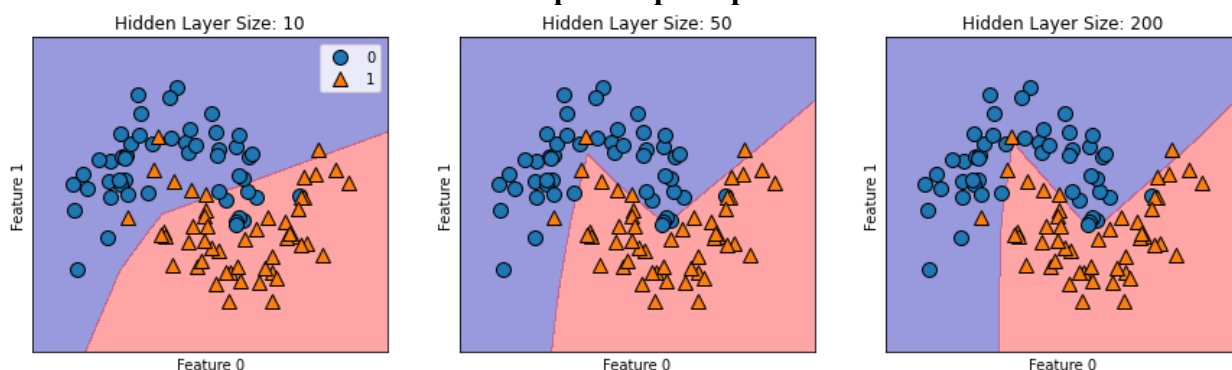


Источник: составлено авторами

Нейронные сети (Многослойный персептрон). Нейронная сеть – это модель, вдохновленная структурой человеческого мозга. Основная структура алгоритма состоит из входных, выходных и скрытых слоев. Размер скрытых слоев – важный гиперпараметр алгоритма. На каждом уровне есть узлы, содержащие номер, и каждый узел получает сигнал от каждого узла предыдущего уровня и отправляет сигнал узлам следующего уровня. У каждого сигнала есть вес и смещение, на которое вход не влияет [14].

Рисунок 8

Классификатор нейронной сети



Источник: составлено авторами

Одной из важных особенностей нейронной сети является алгоритм градиентного спуска. Он используется для минимизации отклонения между выводом и расчетным значением вывода. Нейронная сеть использует обратное распространение для вычисления градиентов. Затем он обновляет все смещения и корректирует веса всех узлов, начиная с вывода [14].

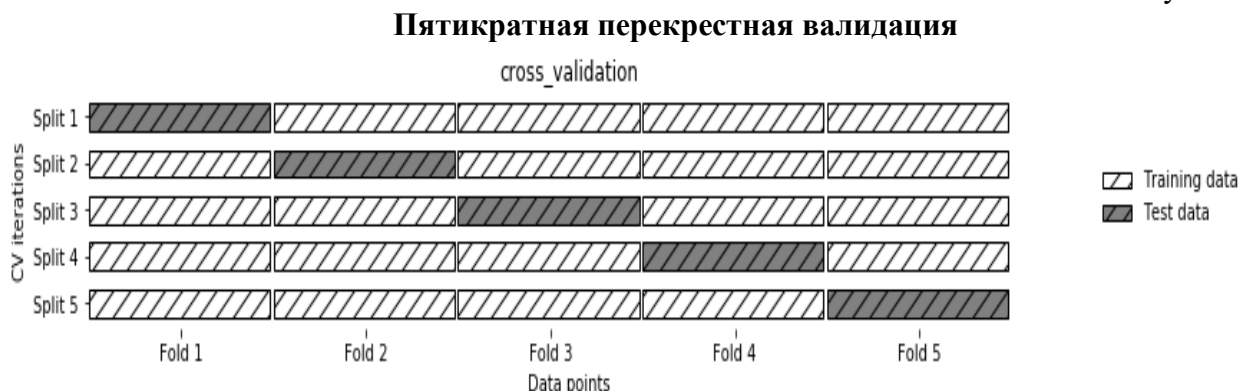
XGBoost. Экстремальный бустинг градиента (XGBoost) – один из широко используемых алгоритмов. Это реализация деревьев решений с градиентным усилением. Скорость выполнения и производительность модели являются преимуществами перед другими алгоритмами повышения градиента [5].

Преимущество метода – масштабируемость. Он использует новый алгоритм изучения дерева для обработки разреженных данных и может автоматически обрабатывать отсутствующие значения. Еще одна важная особенность метода – возможность обрезки деревьев для более глубоких, оптимизированных деревьев. Параллельная и распределенная обработка данных делает его одним из самых быстрых алгоритмов. Алгоритм использует вычисления вне оперативной памяти, что облегчает и ускоряет обработку данных. Следовательно, данный метод является наиболее оптимальным для обработки больших объемов данных. [9].

Кросс-валидация (Cross-validation). Перекрестная проверка – это статистический метод, используемый для измерения производительности методов на основе набора данных обучения и тестирования. Часто используемой версией перекрестной проверки

является перекрестная проверка в k-кратном размере, где количество складок составляет 5 или 10 [19]. В исследовании использовалась 5-кратная перекрестная проверка, как показано на рисунке 9. Следовательно, набор данных делится на пять равных частей. Если первая часть данных используется в качестве тестового набора, остальные являются обучающей выборкой. То есть цель состоит в том, чтобы проанализировать влияние наборов данных на точность модели на основе метрик.

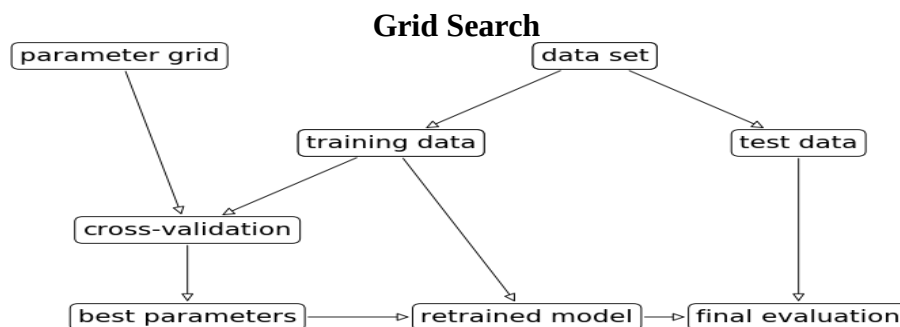
Рисунок 9



Источник: Mglern library (Mueller & Guido, 2017)

Перебор по сетке (Grid Search). Перебор по сетке – это заключительный этап построения модели. На этом этапе настраивается выбранная модель машинного обучения. Сетка параметров – это первый шаг процесса настройки, который включает набор всех возможных гиперпараметров алгоритмов. Затем модель с другим набором гиперпараметров применяется к обучающим и тестовым данным. Этот шаг помогает определить набор параметров, по которым модель получает наивысший показатель. Модель будет реконструирована с использованием набора лучших гиперпараметров [19].

Рисунок 10



Источник: Mglern library (Mueller & Guido, 2017)

Выбор гиперпараметров для настройки модели является значимым вопросом на данном этапе. Учитывая, что перебор параметров занимает много времени, сетка параметров была построена на основе наиболее важных параметров. Например, единственное количество ближайших соседей было параметром, выбранным для настройки модели kNN. Также были использованы сетки параметров, используемые на Kaggle (веб-сайт для специалистов по обработке данных).

Метрики (Metrics). В машинном обучении алгоритмы регрессии, кластеризации и классификации используют разные показатели производительности. Алгоритмы классификации также имеют разные метрики для двоичной и мультиклассовой классификации. В данном исследовании были использованы 6 показателей эффективности моделей:

1. Правильность классификации (Accuracy score);
2. Точность (Precision score);
3. Полнота (Recall score);
4. F-мера (F1 score);
5. Мера Жаккара (Jaccard score);
6. Площадь под кривой ошибок (The area under the receiving operating characteristic (ROC) curve);
7. Доля ошибок второго рода (Type 2 error percentage).

Матрица ошибок (Confusion matrix). Матрица ошибок – это матричная таблица, используемая для оценки эффективности классификатора. Обычно он указывает количество правильных и неправильных результатов алгоритма классификации. Он также предоставляет информацию об ошибках первого и второго рода, которые будут объяснены в этом разделе.

Таблица 2

Confusion matrix		
Predicted class (expectation)	Actual class (observation)	
	TP (true positive) Correct result	FP (false positive) Unexpected result
	FN (false negative) Missing result	TN (true negative) Correct absence of result

Источник: Binary classification (du Boisberranger et al., n.d.)

1. **Правильность классификации (Accuracy score).** Accuracy – самый важный показатель в исследовании. Он также используется в качестве показателя оценки для настройки модели при поиске по сетке. Метрика указывает долю правильных результатов модели:

$$Accuracy(y_{true}, y_{pred}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} 1(if \ y_{true} = y_{pred}) [10]$$

$$= \frac{TP + TN}{TP + FP + FN + TN} [12]; [19]$$

2. **Точность (Precision score).** Показатель точности – это количество правильных положительных прогнозов, деленное на количество общих положительных прогнозов.

$$Precision = \frac{TP}{TP + FP} [12]; [19]$$

3. **Полнота (Recall score).** Полнота – это количество правильных положительных прогнозов, деленное на количество реальных положительных результатов.

$$Recall = \frac{TP}{TP + FN} [12]; [19]$$

4. **F-мера (F1 score).** Точность и полнота – важные показатели. Однако ни один из них по отдельности не даст полной картины. F-мера – еще одна метрика, используемая для двоичной классификации, и это гармоническое среднее значение точности и полноты (Geron, 2019; Mueller & Guido; 2017).

$$F_1 = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} = 2 \times \frac{precision \times recall}{precision + recall} = \frac{TP}{TP + \frac{FN + FP}{2}} [12]$$

5. *Мера Жаккара (Jaccard similarity coefficient score)*. Мера Жаккара (JSC) – это пересечение фактического и прогнозируемого выхода для их объединения.

$$JSC = \frac{TP}{TP + FP + FN} = \frac{F_1}{2 - F_1} [16]$$

6. *Площадь под кривой ошибок (Area under ROC curve)*. Кривая ROC является важным инструментом для анализа производительности двоичных классификаторов. Это линейная кривая, которая отображает процентную ставку TP относительно ставки FP. Площадь под этой кривой – еще один способ сравнения классификаторов [12].

Значение этих показателей колеблется от 0 до 1: чем выше значение, тем больше подходит модель для анализа потребительского кредитования. Кроме того, эти метрики имеют очень тесную взаимосвязь.

7. *Доля ошибок второго рода (Type 2 percentage)*. Классификаторы машинного обучения как статистические модели также имеют ошибки первого и второго рода. Ошибка первого рода также известна как ложноположительный результат (FP) [21]. Например, один из коммерческих банков второго уровня внедряет модели кредитного скоринга. Модель отклоняет выдачу ссуды потребителю, поскольку заявляет, что в будущем она станет неработающей ссудой. Однако на самом деле потребитель способен полностью погасить ссуду. В этой ситуации банки отказываются выпускать долг, но это не окажет существенного влияния на ликвидность и платежеспособность банка.

Ложноотрицательные (FN) – это еще один тип ошибок, который также известен как ошибка второго рода [21]. Предположим, что банк выдает долг потребителю на основе результатов машинного обучения, в которых утверждается, что потребитель обязательно выплатит его. В действительности ссуда становится неработающей, что отрицательно сказывается на ликвидности банка. Более высокий процент ошибки 2-го рода показывает, насколько хуже модель.

4. Обсуждение полученных результатов

Как обсуждалось ранее, методы случайного удаления мажоритарного класса и случайного дублирования миноритарного класса – не лучшая стратегия семплинга. Первая приводит к потере информации, вторая – к проблемам переобучения. Кроме того, случайное удаление мажоритарного класса было единственным вариантом адаптации нормативных данных в данном исследовании. В этом разделе будет обсуждаться результат всех классификаторов на основе данных, к которым они были применены.

Субдискредитированные данные (Undersampling data). Результаты исследования, линейные классификаторы и наивный байесовский классификатор не являются лучшими вариантами моделирования кредитного скоринга. Таблица 3 демонстрирует, что у всех классификаторов были проблемы с недостаточным соответствием: алгоритмы плохо моделировали обучающие данные и некорректно выполнили задачу на данных тестирования.

Таблица 3

Accuracy of classifiers applied to undersampled data

	Linear Models		Non-Linear Models					
	Logistic Regression	SGD	Naïve Bayes	kNN	Decision Tree	Random Forest	Neural Networks	XGB
Training	58,6%	58,7%	59,0%	68,6%	68,3%	64,9%	70,8%	69,6%
Testing	58,7%	58,9%	59,0%	67,1%	65,9%	64,6%	70,0%	67,2%

Источник: составлено авторами

Несмотря на низкие показатели по другим метрикам, классификатор стохастического градиентного спуска (SGD) показал одну из самых низких долей ошибок второго рода. В этом время классификатор нейронных сетей показал лучший результат среди моделей, примененных к данным с недостаточной выборкой на основе большинства метрик. Однако высокий процент ошибок 2-го рода, а также необходимость большого количества времени для обработки данных не делает его фаворитом. В итоге модель классификатора экстремального бустинга градиента (XGB) является наиболее подходящим вариантом моделирования кредитного скоринга с оптимальными показателями метрик: второй по величине показатель правильности классификации и второй по величине процент ошибок второго рода.

Таблица 4

Models and performance results

		Metrics						
		Accuracy	Precision	Recall	F1	JSC	AUC_ROC	Type 2 error
Linear	Logistic Regression	58,7%	59,0%	58,5%	58,7%	41,6%	58,7%	20,8%
	SGD	58,9%	57,9%	66,5%	61,9%	44,8%	58,9%	16,8%
Non-Linear	Naïve Bayes	59,0%	60,1%	54,3%	57,0%	39,9%	59,0%	23,0%
	kNN	67,1%	68,4%	64,0%	66,1%	49,4%	67,1%	18,6%
	Decision Tree	65,9%	66,9%	63,5%	65,2%	48,3%	65,9%	18,3%
	Random Forest	64,6%	63,5%	69,4%	66,3%	49,6%	64,6%	15,4%
	Neural Networks	70,0%	72,8%	64,1%	68,2%	51,7%	70,0%	18,0%
	XGB	67,2%	67,4%	67,3%	67,4%	50,8%	67,2%	16,4%

Источник: составлено авторами

Передискредитированные данные (Oversampled data). Классификаторы в целом намного лучше работали с передискредитированными данными, сформированными на основе метода повторной выборки SMOTE, который привнес в данные дополнительную информацию. При этом наихудшие результаты показали линейные классификаторы и наивный байесовский классификатор. Другие классификаторы показали положительные результаты при работе с набором данных для обучения, а также передискредитированными данными.

Таблица 5

Accuracy of classifiers applied to oversampled data

	Linear Models		Non-Linear Models					
	Logistic Regression	SGD	Naïve Bayes	kNN	Decision Tree	Random Forest	Neural Networks	XGB
Training	64,1%	64,1%	64,9%	99,9%	76,8%	99,6%	73,6%	74,6%
Testing	64,1%	64,1%	64,9%	83,5%	75,3%	84,8%	73,6%	74,4%

Источник: составлено авторами

Нейронные сети были многообещающей моделью, но есть незначительная разница между показателем модели для субдискредитированных и передискредитированных данных. Несмотря на повышение точности передискредитированных данных по сравнению с субдискредитированными данными, нейронные сети не смогли превзойти нелинейные модели.

Согласно таблице 6, классификаторы k-ближайших соседей и случайных лесов превосходили другие классификаторы по всем показателям. Кроме того, обе модели продемонстрировали самый низкий процент ошибок 2-го рода. Однако классификатор

случайного леса превзошел классификатор k-ближайших соседей по всем показателям, кроме точности.

Таблица 6

Models and performance results

		Metrics						
		Accuracy	Precision	Recall	F1	JSC	AUC_ROC	Type 2 error
Linear	Logistic Regression	64,1%	63,9%	65,1%	64,5%	47,6%	64,1%	17,5%
	SGD	64,1%	63,9%	65,1%	64,5%	47,6%	64,1%	17,5%
Non-Linear	Naïve Bayes	64,9%	62,6%	74,2%	67,9%	51,4%	64,9%	12,9%
	kNN	83,5%	85,2%	81,0%	83,0%	71,0%	83,5%	9,5%
	Decision Tree	75,3%	74,8%	76,2%	75,5%	60,7%	75,3%	11,9%
	Random Forest	84,8%	84,7%	85,0%	84,8%	73,6%	84,8%	7,5%
	Neural Networks	73,6%	71,2%	79,1%	75,0%	60,0%	73,6%	10,5%
	XGB	74,4%	73,5%	76,2%	74,9%	59,8%	74,4%	11,9%

Источник: составлено авторами

Заключение

Результаты исследования показали, что модели машинного обучения достаточно хорошо работают на основе регуляторных данных, собранных центральным банком. Кроме того, анализ потребительских кредитов с помощью алгоритмов машинного обучения продемонстрировал уникальное понимание особенностей (плохих и хороших) потребительских заемщиков, а также предоставил информацию для проверки правильности выданных потребительских кредитов банками.

В этом исследовании мы показали, что очень важно проверять качество данных (во время процедур упорядочения и очистки, чтобы исключить ненужные переменные) и работать с несбалансированным набором обучающих данных, чтобы избежать смещения в пользу большинства категорий.

В части оценки точности прогнозов моделей передискредитированные данные, скорректированные методом SMOTE, показали более многообещающие результаты по сравнению с данными, обработанными стратегией случайного удаления мажоритарного класса. Другими словами, передискредитированные данные, скорректированные с помощью SMOTE, помогли минимизировать потерю информации и повысить эффективность прогнозирования. Модели с хорошо подобранными обучающими данными лучше работали с передискредитированными данными по сравнению с субдискредитированными данными.

Кроме того, нелинейные модели иллюстрировали более точные прогнозы по сравнению с линейными моделями. В частности, нелинейные модели, такие как классификаторы случайного леса и k-ближайших соседей на данных, превзошли другие модели классификаторов. С другой стороны, линейные модели, такие как логистическая регрессия и классификатор SGD, показали наихудшие результаты среди восьми сравниваемых моделей.

Подводя итог, можно сделать вывод, что модели, основанные на регуляторных данных, могут стать адекватной основой для оценки кредитного риска по выданным потребительским кредитам банками второго уровня, а также могут помочь центральному банку прогнозировать потенциальные систематические риски. Итак, согласно результатам исследования, оценка кредитного риска с помощью машинного обучения может стать хорошим дополнением к регулированию банков второго уровня, выдающих потребительские кредиты.

Для дальнейшего развития данного исследования предполагается несколько направлений. В первую очередь необходимо дополнить имеющиеся наборы данных

социально-демографическими характеристиками заемщиков, которые могут положительно повлиять на производительность алгоритмов [13].

В части моделирования предполагается несколько подходов, которые могут повысить эффективность моделей:

- а) применение гибридных методов машинного обучения;
- б) применение более современных алгоритмов машинного обучения (CatBoost, Light GBM);
- с) включение дополнительных параметров в сетку параметров и стратегии сэмплинга.

Список использованной литературы

1. Addo, P., Guegan, D., & Hassani, B. (2018). Credit Risk Analysis Using Machine and Deep Learning Models. *Risks*, 6(2), 38. doi: 10.3390/risks6020038.
2. Alencar, R. (2017). Resampling strategies for imbalanced datasets. Retrieved from <https://www.kaggle.com/rafjaa/resampling-strategies-for-imbalanced-datasets>.
3. Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of The Operational Research Society*, 54(6), 627-635. doi: 10.1057/palgrave.jors.2601545.
4. Brown, I., & Mues, C. (2012). An experimental comparison of classification algorithms for imbalanced credit scoring data sets. *Expert Systems with Applications*, 39(3), 3446-3453. doi: 10.1016/j.eswa.2011.09.033.
5. Brownlee, J. (2016). A Gentle Introduction to XGBoost for Applied Machine Learning. Retrieved from <https://machinelearningmastery.com/gentle-introduction-xgboost-applied-machine-learning/>.
6. Brownlee, J. (2016). Metrics to Evaluate Machine Learning Algorithms in Python. Retrieved from <https://machinelearningmastery.com/metrics-evaluate-machine-learning-algorithms-python/>.
7. Brownlee, J. (2020). SMOTE for Imbalanced Classification with Python. Retrieved from <https://machinelearningmastery.com/smote-oversampling-for-imbalanced-classification/>.
8. Chawla, N., Bowyer, K., Hall, L., & Kegelmeyer, W. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321-357. Doi: 10.1613/jair.953.
9. Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of The 22Nd ACM SIGKDD International Conference On Knowledge Discovery and Data Mining*, 785–794. doi: 10.1145/2939672.2939785.
10. du Boisberranger, J., Van den Bossche, J., Esteve, L., Fan, T., Gramfort, A., & Grisel, O. et al. User guide: contents – scikit-learn 0.23.2 documentation. Retrieved from https://scikit-learn.org/stable/user_guide.html.
11. Fuchs, M. (2019). Introduction to SGD Classifier - Michael Fuchs Python. Retrieved from <https://michael-fuchs-python.netlify.app/2019/11/11/introduction-to-sgd-classifier/>.
12. Geron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). Sebastopol, California: O'Reilly Media, Incorporated.
13. Grier, W. (2012). Credit analysis of financial institutions (2nd ed., pp. 294-296). London: Euromoney Institutional Investor PLC.
14. Hansen, C. (2019). Neural Networks: Feedforward and Backpropagation Explained. Retrieved from <https://mlfromscratch.com/neural-networks-explained/#overview>.
15. Henley, W., & Hand, D. (1996). A k-Nearest-Neighbour Classifier for Assessing Consumer Credit Risk. *The Statistician*, 45(1), 77. doi: 10.2307/2348414.

16. Labetut, V., & Cherifi, H. (2011). Evaluation of Performance Measures for Classifiers Comparison. *Ubiquitous Computing and Communication Journal*, 6, 21-34. Retrieved from <https://arxiv.org/abs/1112.4133>.
17. Lemaitre, G., Nogueira, F., Oliveira, D., & Aridas, C. 2. Over-sampling – imbalanced-learn 0.5.0 documentation. Retrieved from https://imbalanced-learn.readthedocs.io/en/stable/over_sampling.html#smote-adasyn.
18. Lessmann, S., Baesens, B., Seow, H., & Thomas, L. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, 247(1), 124-136. doi: 10.1016/j.ejor.2015.05.030.
19. Mueller, A., & Guido, S. (2017). *Introduction to Machine Learning with Python* (1st ed.). Sebastopol, California: O'Reilly Media.
20. Munkhdalai, L., Munkhdalai, T., Namsrai, O., Lee, J., & Ryu, K. (2019). An Empirical Comparison of Machine-Learning Methods on Bank Client Credit Assessments. *Sustainability*, 11(3), 699. doi: 10.3390/su11030699.
21. Schmarzo, B. (2018). Understanding Type 1 and Type 2 Errors [Blog]. Retrieved from <https://www.datasciencecentral.com/profiles/blogs/understanding-type-i-and-type-ii-errors>
22. Tsai, C., & Chen, M. (2010). Credit rating by hybrid machine learning techniques. *Applied Soft Computing*, 10(2), 374-380. doi: 10.1016/j.asoc.2009.08.003.

Приложение

Важные библиотеки

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib
import os
path = 'E:\...' # выбор пути к документу, в котором находится файл
os.chdir(path)
df_ml = pd.read_csv('filename.csv')
X_df = df_ml.iloc[:, :-1].values
y_df = df_ml.iloc[:, -1].values
```

Случайное удаление мажоритарного класса с библиотеки imblearn

```
from imblearn.under_sampling import RandomUnderSampler as rus
us = rus(random_state=42)
X, y = us.fit_resample(X_df, y_df)
```

Стратегия SMOTE-NC с библиотеки imblearn

```
from imblearn.over_sampling import SMOTENC
sm = SMOTENC(random_state=42, categorical_features=[0,1,2,3,4,5]) # в исходных данных
# имеются категориальные переменные
X, y = sm.fit(X_df, y_df)
```

Кодирование категориальных переменных

```
from sklearn.compose import ColumnTransformer
from sklearn.preprocessing import OneHotEncoder
```

первый столбец категориальных переменных содержал 18 различных переменных (три города и региона (также бывшее название для одного региона)), таким образом создавалось 18 строк

```
ct0 = ColumnTransformer(transformers=[('encoder', OneHotEncoder(), [0])],
remainder='passthrough')
X=np.array(ct0.fit_transform(X))
```

второй столбец категориальных переменных содержал 4 разных переменных (четыре типа валюты, в которой был выдан заем), таким образом создавая 4 строки

```
ct1 = ColumnTransformer(transformers=[('encoder', OneHotEncoder(), [18])],
remainder='passthrough')
X=np.array(ct1.fit_transform(X))
```

третий столбец категориальных переменных содержал 2 разные переменные (кредитная карта или заем), таким образом создавая 2 строки

```
ct2 = ColumnTransformer(transformers=[('encoder', OneHotEncoder(), [18])],
remainder='passthrough')
X=np.array(ct2.fit_transform(X))
```

четвертый столбец категориальных переменных содержал 2 разные переменные (потребительское или автокредитование), таким образом создавая 2 строки

```
ct3 = ColumnTransformer(transformers=[('encoder', OneHotEncoder(), [18])],
remainder='passthrough')
```

```
X=np.array(ct3.fit_transform(X))
```

```
# пятый столбец категориальных переменных содержал 2 разные переменные (мужской  
или женский), таким образом создавая 2 строки
```

```
ct4 = ColumnTransformer(transformers=[('encoder', OneHotEncoder(),[18])],
```

```
remainder='passthrough')
```

```
X=np.array(ct4.fit_transform(X))
```

```
# шестой столбец категориальных переменных содержал 2 разные переменные (резидент  
или нерезидент), таким образом создавая 2 строки
```

```
ct5 = ColumnTransformer(transformers=[('encoder', OneHotEncoder(),[18])],
```

```
remainder='passthrough')
```

```
X=np.array(ct5.fit_transform(X))
```

```
# Label Encoder переменной y
```

```
from sklearn.preprocessing import LabelEncoder
```

```
le = LabelEncoder()
```

```
y = le.fit_transform(y)
```

```
# разделение данных на обучающий и тестовый набор данных
```

```
from sklearn.model_selection import train_test_split
```

```
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.2, random_state=1)
```

```
# шкалирование
```

```
from sklearn.preprocessing import StandardScaler
```

```
sc = StandardScaler()
```

```
X_train[:, -3:] = sc.fit_transform(X_train[:, -3:])
```

```
X_test[:, -3:] = sc.transform(X_test[:, -3:])
```

```
# Метод главных компонент (Principal component analysis (PCA))
```

```
from sklearn.decomposition import PCA
```

```
pca = PCA(n_components=a) # a is quantity where cumulative explained variance ratio > 95%
```

```
X_train = pca.fit_transform(X_train)
```

```
X_test = pca.transform(X_test)
```

```
# Логистическая регрессия
```

```
from sklearn.linear_model import LogisticRegression
```

```
log = LogisticRegression().fit(X_train,y_train)
```

```
y_tr_log_pred = log.predict(X_train)
```

```
y_ts_log_pred = log.predict(X_test)
```

```
# Стохастическая градиентный спуск
```

```
from sklearn.linear_model import SGDClassifier
```

```
sgd=SGDClassifier().fit(X_train,y_train)
```

```
y_tr_sgd_pred = sgd.predict(X_train)
```

```
y_ts_sgd_pred = sgd.predict(X_test)
```

```
# Гауссовский наивный байесовский классификатор
```

```
from sklearn.naive_bayes import GaussianNB
```

```
nbc = GaussianNB().fit(X_train,y_train)
```

```
y_tr_nb_pred = nbc.predict(X_train)
```

```
y_ts_nb_pred = nbc.predict(X_test)
```

k-ближайшие соседи

```
from sklearn.neighbors import KNeighborsClassifier
knn=KNeighborsClassifier().fit(X_train,y_train)
y_tr_knn_pred = knn.predict(X_train)
y_ts_knn_pred = knn.predict(X_test)
```

Дерево решений

```
from sklearn.tree import DecisionTreeClassifier
dtc=DecisionTreeClassifier(max_depth=14, criterion='entropy').fit(X_train,y_train)
y_tr_dt_pred = dtc.predict(X_train)
y_ts_dt_pred = dtc.predict(X_test)
```

Случайный лес

```
from sklearn.ensemble import RandomForestClassifier
rfc=RandomForestClassifier().fit(X_train,y_train)
y_tr_rf_pred = rfc.predict(X_train)
y_ts_rf_pred = rfc.predict(X_test)
```

Многослойный персептрон

```
from sklearn.neural_network import MLPClassifier
nnc = MLPClassifier().fit(X_train,y_train)
y_tr_nnc_pred = nnc.predict(X_train)
y_ts_nnc_pred = nnc.predict(X_test)
```

XGBoost

```
from xgboost import XGBClassifier
xgb = XGBClassifier().fit(X_train,y_train)
y_tr_xgb_pred = xgb.predict(X_train)
y_ts_xgb_pred = xgb.predict(X_test)
```

Кросс-валидация

```
from sklearn.model_selection import cross_val_score
accuracies = cross_val_score(estimator = xgb, X = X_train, y = y_train, cv = 5)
print('Accuracy: {:.2f} %'.format(accuracies.mean()*100))
print('Standard deviation: {:.2f} %'.format(accuracies.std()*100))
```

Примечание: мы можем поставить другие модели вместо XGB для анализа результата перекрестной проверки.

Перебор по сетке для логистической регрессии

```
from sklearn.model_selection import GridSearchCV
parameters = [{'penalty': ['none'], 'solver':['newton-cg', 'sag', 'saga', 'lbfgs']},
               {'penalty': ['elasticnet'], 'C': [0.01, 0.1, 0.25, 0.5, 0.75, 1, 5, 10], 'solver':['saga']},
               {'penalty': ['l2'], 'C': [0.01, 0.1, 0.25, 0.5, 0.75, 1, 5, 10], 'solver':['newton-cg', 'sag',
'saga', 'lbfgs']}]
grid_search = GridSearchCV(estimator = log,
                           param_grid = parameters,
                           scoring = 'accuracy',
                           cv = 5,
                           n_jobs = -1)
grid_search.fit(X_train, y_train)
```

```
best_accuracy = grid_search.best_score_  
best_parameters = grid_search.best_params_  
print('Best accuracy: {:.2f} %'.format(best_accuracy*100))  
print('Best parameters: ',best_parameters)
```

Перебор по сетке для стохастического градиента спуска

```
from sklearn.model_selection import GridSearchCV  
parameters = [{"loss": ["hinge", "log", "squared_hinge", "modified_huber"],  
               "alpha": [0.0001, 0.001, 0.01, 0.1], "penalty": ["l2", "l1", "none"]}]  
grid_search = GridSearchCV(estimator = sgd,  
                           param_grid = parameters,  
                           scoring = 'accuracy',  
                           cv = 5,  
                           n_jobs = -1)  
grid_search.fit(X_train, y_train)  
best_accuracy = grid_search.best_score_  
best_parameters = grid_search.best_params_  
print('Best accuracy: {:.2f} %'.format(best_accuracy*100))  
print('Best parameters: ',best_parameters)
```

Перебор по сетке для Гауссовскому наивному байесовскому классификатору

```
from sklearn.model_selection import GridSearchCV  
parameters = [{'var_smoothing': [1e-12,1e-10,1e-7,1e-4,1e-3,1e-2,1e-1,1,10]}]  
grid_search = GridSearchCV(estimator = nbc,  
                           param_grid = parameters,  
                           scoring = 'accuracy',  
                           cv = 5,  
                           n_jobs = -1)  
grid_search.fit(X_train, y_train)  
best_accuracy = grid_search.best_score_  
best_parameters = grid_search.best_params_  
print('Best accuracy: {:.2f} %'.format(best_accuracy*100))  
print('Best parameters: ',best_parameters)
```

Перебор по сетке для k-ближайших соседей

```
from sklearn.model_selection import GridSearchCV  
parameters = [{'n_neighbors': list(range(1, 81))}]  
grid_search = GridSearchCV(estimator = knn,  
                           param_grid = parameters,  
                           scoring = 'accuracy',  
                           cv = 5,  
                           n_jobs = -1)  
grid_search.fit(X_train, y_train)  
best_accuracy = grid_search.best_score_  
best_parameters = grid_search.best_params_  
print('Best accuracy: {:.2f} %'.format(best_accuracy*100))  
print('Best parameters: ',best_parameters)
```

Перебор по сетке для дерева решений

```
from sklearn.model_selection import GridSearchCV  
parameters = [{'criterion':['gini', 'entropy'],'max_depth': list(range(1,21))}]  
grid_search = GridSearchCV(estimator = dtc,
```

```
        param_grid = parameters,
        scoring = 'accuracy',
        cv = 5,
        n_jobs = -1)
grid_search.fit(X_train, y_train)
best_accuracy = grid_search.best_score_
best_parameters = grid_search.best_params_
print('Best accuracy: {:.2f} %'.format(best_accuracy*100))
print('Best parameters: ',best_parameters)
```

Перебор по сетке для случайного леса

```
from sklearn.model_selection import GridSearchCV
parameters = [{'n_estimators': list(range(1,21)), 'max_features': ['auto', 'sqrt', 'log2'],
               'max_depth' : ['None',8], 'criterion' :['gini', 'entropy']}]
grid_search = GridSearchCV(estimator = rfc,
                           param_grid = parameters,
                           scoring = 'accuracy',
                           cv = 5,
                           n_jobs = -1)
grid_search.fit(X_train, y_train)
best_accuracy = grid_search.best_score_
best_parameters = grid_search.best_params_
print('Best accuracy: {:.2f} %'.format(best_accuracy*100))
print('Best parameters: ',best_parameters)
```

Перебор по сетке для многослойного персептрона

```
from sklearn.model_selection import GridSearchCV
parameters = [{'hidden_layer_sizes':[100,200,300,[200,50],[100,100],[200,100]],
               'activation':['identity','logistic','tanh','relu'],
               'solver': ['adam'],
               'learning_rate':['constant','invscaling','adaptive'],
               'max_iter': [1000,1500,2000 ]}]
grid_search = GridSearchCV(estimator = nnc,
                           param_grid = parameters,
                           scoring = 'accuracy',
                           cv = 5,
                           n_jobs = -1)
grid_search.fit(X_train, y_train)
best_accuracy = grid_search.best_score_
best_parameters = grid_search.best_params_
print('Best accuracy: {:.2f} %'.format(best_accuracy*100))
print('Best parameters: ',best_parameters)
```

Перебор по сетке для XGBoost

```
from sklearn.model_selection import GridSearchCV
parameters = [{'n_estimators': [1000], #number of trees, change it to 1000 for better results
               'max_depth': [6,7,8],
               'learning_rate': [0.05], #so called `eta` value
               'objective':['binary:logistic'],
               'tree_method':['exact'],
               'min_child_weight': [11],
               'subsample': [0.8],
```

```
'colsample_bytree': [0.7]})
grid_search = GridSearchCV(estimator = xgb,
                           param_grid = parameters,
                           scoring = 'accuracy',
                           cv = 5,
                           n_jobs = -1)
grid_search.fit(X_train, y_train)
best_accuracy = grid_search.best_score_
best_parameters = grid_search.best_params_
print('Best accuracy: {:.2f} %'.format(best_accuracy*100))
print('Best parameters: ',best_parameters)
```

После перебора по сетке должны быть применены все лучшие параметры, чтобы быть уверенным, что модель с наилучшей точностью дает более высокую точность, и после этого следует проанализировать другие метрики.

Метрики

```
from sklearn.metrics import confusion_matrix, accuracy_score
knn_cm_tr = confusion_matrix(y_train,y_tr_knn_pred)
print(knn_cm_tr)
accuracy_score(y_train, y_tr_knn_pred)
```

```
knn_cm_ts = confusion_matrix(y_test,y_ts_knn_pred)
print(knn_cm_ts)
accuracy_score(y_test, y_ts_knn_pred)
```

```
from sklearn.metrics import roc_auc_score, jaccard_score, f1_score, precision_score,
recall_score
print(roc_auc_score(y_test,y_ts_knn_pred))
print(jaccard_score(y_test,y_ts_knn_pred))
print(f1_score(y_test,y_ts_knn_pred))
print(precision_score(y_test,y_ts_knn_pred))
print(recall_score(y_test,y_ts_knn_pred))
```

Здесь метрики использовались для анализа производительности классификатора k-ближайших соседей, переменные должны быть изменены для достижения результата других моделей.

ЛИКВИДНОСТЬ РЫНКА ГЦБ: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ РЕШЕНИЙ

Алимбетова И.С. – главный специалист-аналитик управления анализа рынка финансовых и реальных активов департамента финансовой стабильности и исследований Национального Банка Республики Казахстан.

Кадырбеков Т.Ж. – ведущий специалист-аналитик управления анализа рынка финансовых и реальных активов департамента финансовой стабильности и исследований Национального Банка Республики Казахстан.

Мустафин Е.Т. – заместитель директора департамента финансовой стабильности и исследований Национального Банка Республики Казахстан.

Ыбраев Ж.Ж. – начальник управления анализа рынка финансовых и реальных активов департамента финансовой стабильности и исследований Национального Банка Республики Казахстан.

В этой статье обсуждаются факторы, лежащие в основе низкой ликвидности и глубины рынка государственных ценных бумаг Казахстана. Рассматриваются потенциал развития рынка, варианты улучшения его структуры и функционирования. Мы концентрируемся на микроструктурных характеристиках рынка ГЦБ, влияющих на уровень его ликвидности. Обсуждаем факторы, такие как структура долговых инструментов, конкуренция и концентрация, управление номенклатурой, а также уровень прозрачности информации. В частности, мы предлагаем обсудить и попытаться понять, как те или иные изменения в некоторых из этих факторов с течением времени могут повлиять на эволюцию ликвидности, глубину и структурированность рынка.

Ключевые слова: государственные ценные бумаги, ликвидность, государственный долг, рынок ценных бумаг.

JEL-классификация: H63, G11, G14, G15.

1. Введение

Во многих странах рынок государственного долга является важнейшим звеном в эффективном функционировании всей финансовой системы. Поскольку рынок ГЦБ выполняет несколько ключевых функций, центральные банки, практикующие политику инфляционного таргетирования, проявляют к нему особый интерес. Именно здесь они часто проводят свои внутренние денежно-кредитные операции и отсюда они извлекают информацию об ожиданиях будущих изменений процентных ставок. На этом рынке правительства привлекают средства, а для центральных банков, на которых возложена функция агента размещения, этот рынок становится основным пространством проведения своей политики. В силу своего практически безрискового характера государственные ценные бумаги служат ориентиром ценообразования и инструментом хеджирования для других ценных бумаг с фиксированным доходом. Они служат в качестве обеспечения или как часть нормативных требований для различных финансовых посредников, позволяя им финансировать свои операции. Наконец, поскольку другие рынки с фиксированным доходом обладают многими структурными и институциональными характеристиками рынка ГЦБ, более глубокое понимание того, как функционирует рынок государственных ценных бумаг, позволяет центральным банкам лучше понимать другие рынки с фиксированным доходом. Очевидно, что ликвидность является важным аспектом всех финансовых рынков, и задача повышения и поддержания ликвидности рынка ГЦБ должна стать одной из важнейших как для центрального банка, так и для правительства [10].

Из-за своей многомерной природы концепция рыночной ликвидности не поддается простому определению или легкому измерению. Хотя большинство участников рынка в

некоторой степени определяет, является ли конкретный рынок ликвидным, немногие смогут точно определить все факторы, влияющие на ликвидность этого рынка. Тем не менее существует определенный консенсус: ликвидный рынок – это такой рынок, на котором крупные сделки заключаются быстро, практически не влияя на цены [2].

В литературе рыночная ликвидность часто определяется по четырем аспектам: оперативность, глубина, ширина и устойчивость. Под оперативностью понимается скорость, с которой заключается сделка заданного размера; под глубиной – максимальный размер сделки для заданного спреда; под шириной – затраты на предоставление ликвидности (узкий спред подразумевает большую ликвидность); под устойчивостью – быстрота возвращения цен к их равновесному уровню после крупной сделки или как быстро исчезает дисбаланс при поточных сделках [1]. При этом различные аспекты ликвидности могут сильно взаимодействовать.

Ликвидность рынка ГЦБ влияет на основную деятельность центрального банка тремя способами. Во-первых, ликвидность рынка влияет на деятельность по формированию и реализации денежно-кредитной политики, на эффективность операций на открытом рынке. Так как регулятор на основе данных, извлекаемых из цен, получает информацию о текущих и ожидаемых денежно-кредитных условиях, ликвидность определяет, насколько объективно в ней учитывается сам процесс ценообразования.

Таким образом, степень ликвидности рынка влияет на уверенность центрального банка в его ожиданиях и того, насколько эффективно рынок способен агрегировать в цены информацию об отдельных категориях инвесторов. Низкий уровень рыночной ликвидности препятствует переносу мер денежно-кредитной политики на инструменты с фиксированной доходностью с более длительным сроком доплат.

Во-вторых, в некоторых случаях неликвидность рынка часто является одним из симптомов, если не причиной, системных финансовых сбоев. При низком уровне рыночной ликвидности стрессовые шоки для финансовых рынков усиливаются за счет сопутствующих проблем доверия, оборотной ликвидности и платежеспособности ключевых финансовых игроков. Эти проблемы могут привести к сбоям в работе платежных систем и снизить эффективность распределения капитала. Тогда как ликвидный рынок, наоборот, способствует ослаблению шоков, мотивируя цепную паническую реакцию некоторых участников, так как несет больше информации по ожиданиям в денежно-кредитных условиях. С другой стороны, колебания рыночной ликвидности могут иметь прямое влияние на деятельность надзорных органов как в качестве кредитора последней инстанции, так и монитора финансовой стабильности. Кроме того, при расчете потенциальных рыночных рисков часто игнорируются риски ликвидности и ликвидации, определяемые как риск невозможности своевременно ликвидировать позицию по разумной цене. В частности, методология VaR (Value-at-Risk) предполагает, что цены изменяются непрерывно, и игнорирует возможность того, что движение цен может быть прерывистым в условиях высоких рисков ликвидации [9]. Эти факторы могут вызывать и усугублять сбои в финансовой стабильности.

В-третьих, центральный банк разделяет желание правительства минимизировать затраты на обслуживание долга. Высокая ликвидность вторичного рынка, как правило, способствует выпуску правительством больших объемов долга по относительно низким ставкам. Это связано с тем, что активный спрос инвесторов на первичном рынке напрямую зависит от возможности предсказуемо торговать на вторичном рынке. Таким образом, правительство в роли агента по размещению долга должно работать вместе с центральным банком над повышением целостности и эффективности рынка государственных ценных бумаг.

На ликвидность влияет структура рынка. Обычно на рынках ГЦБ цены формируются котировками множества участников. Следовательно, изучение факторов, влияющих на стимулы участников рынка, является одним из способов оценки уровня ликвидности рынка. Также общая ликвидность рынка зависит от ликвидности конкретных

ценных бумаг. Ликвидность бумаги в свою очередь зависит от нескольких факторов, включая ее объем и эффективное предложение. Другие факторы, влияющие на уровень ликвидности на рынке ГЦБ, включают прозрачность, стоимость транзакции, волатильность процентных ставок и активность на смежных рынках.

Данный обзор состоит из трех частей. Мы концентрируемся на микроструктурных характеристиках рынка ГЦБ, влияющих на его уровень ликвидности. Обсуждаем факторы, такие как структура долговых инструментов, конкуренция и концентрация, управление номенклатурой, а также уровень прозрачности информации. В частности, мы предлагаем обсудить и попытаться понять, как те или иные изменения в некоторых из этих факторов с течением времени могут повлиять на эволюцию ликвидности, глубину и структурированность рынка.

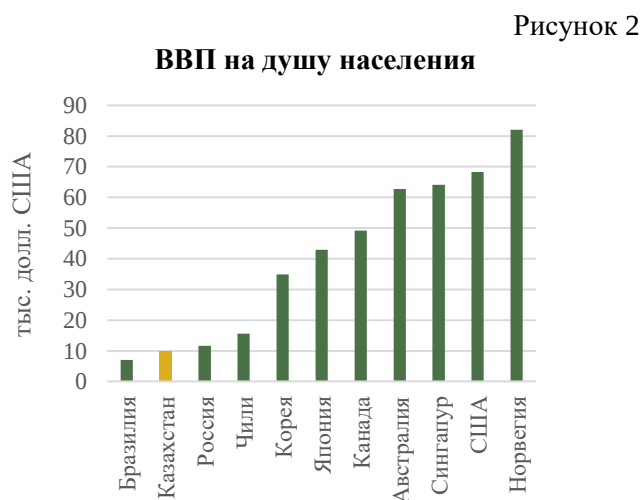
2. Долговой рынок Казахстана

Рынок государственного долга Казахстана появился еще в 1990-е годы и вырос с появлением пенсионной системы [31]. На текущий момент рынок государственного долга имеет достаточно существенный объем, хотя и в значительной степени ниже, чем госдолг по отношению к ВВП в более развитых странах. Несмотря на продолжительную историю, рынок государственных ценных бумаг, как и весь фондовый рынок Казахстана в целом, остается неразвитым, неликвидным и неэффективным. Доверие к фондовому рынку в Казахстане остается низким, и основными финансовыми посредниками выступают банки. В отличие от стран с развитым фондовым рынком финансирование бизнеса в Казахстане в большей степени происходит через банковскую систему, а домашние хозяйства предпочитают хранить средства в банках, а не инвестировать их на бирже. Эти факторы снижают как число эмитентов, так и число инвесторов на фондовом рынке. Основная доля ценных бумаг, имеющих листинг на Казахстанской фондовой бирже, приходится на ГЦБ, ценные бумаги компаний квазигосударственного и финансового секторов.

Рынок институциональных инвесторов на рынке РК, и без того ограниченный, еще больше сузился с объединением пенсионных фондов в 2014 году. Это практически уничтожило ликвидность рынка. При этом для Министерства финансов РК повышение ликвидности рынка ГЦБ не являлось приоритетной задачей, так как, с одной стороны, значительная часть дефицита бюджета финансировалась за счет трансфертов из Национального фонда РК, а, с другой стороны, наличие каптивных инвесторов поддерживало спрос на государственный долг.

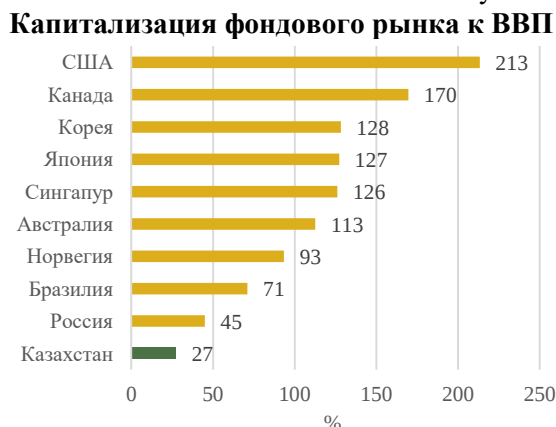


Источник: МВФ, по состоянию на 01.04.2021



Источник: МВФ, по состоянию на 01.04.2021

Рисунок 3



Источник: Bloomberg, KASE.

Примечание: данные по РК приведены по результатам 2020 года, данные по остальным странам приведены по состоянию на 03.06.2021.

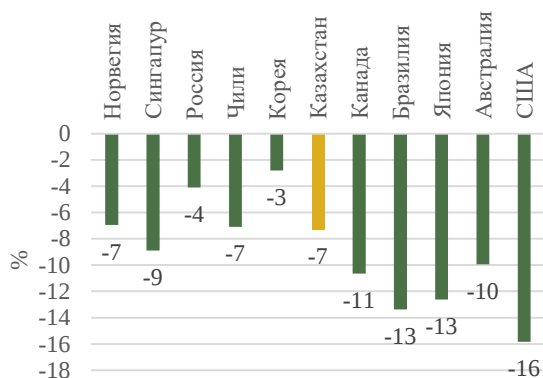
Рисунок 4



Источник: МВФ

Рисунок 5

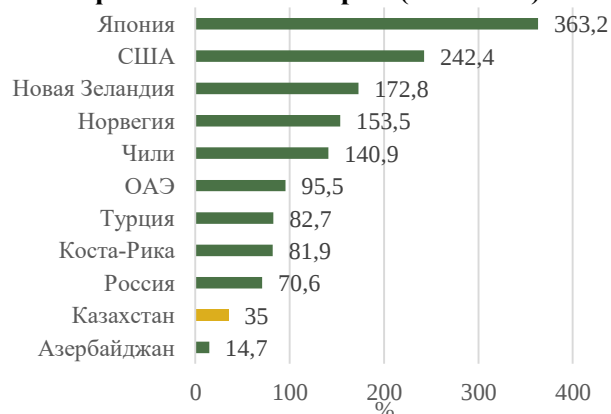
Чистое кредитование/заимствование правительства (% к ВВП)



Источник: МВФ

Рисунок 6

Внутренний кредит, предоставленный финансовым сектором (% к ВВП)



Источник: Всемирный банк

3. Ликвидность и структура рынка ГЦБ

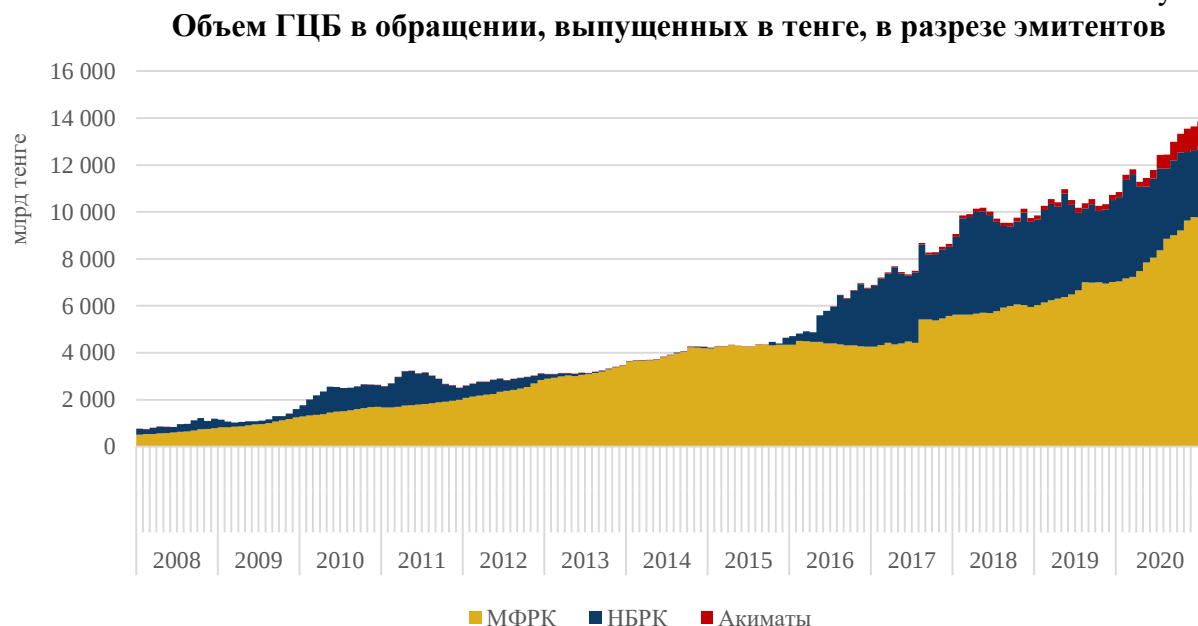
На конец 2020 года общий объем ГЦБ РК в обращении, выпущенных в тенге, составил 14,3 трлн тенге (рисунок 7). Наиболее значительная доля ГЦБ в обращении приходится на казначейские облигации, выпущенные Министерством финансов РК (далее – ГЦБ МФ РК). Так, в 2013-2015 годы основным эмитентом ГЦБ было МФ РК. Однако в 2016 году НБ РК активизировал размещение нот в рамках политики управления процентными ставками, что привело к снижению доли ГЦБ МФ РК в среднем до 60-65% (рисунок 7).

Основной эмитент государственного долга – МФ РК, выпускающее облигации в тенге, а также еврооблигации в иностранной валюте, ГЦБ с фиксированным купоном и ГЦБ с индексированным купоном. Для построения безрисковой кривой доходности в среднем и долгосрочном горизонте ключевую роль играют ГЦБ с фиксированным купоном, выпущенные в тенге. При этом с недавних пор МФ РК возобновило выпуск дисконтных облигаций со сроком до года и ГЦБ с фиксированным купоном со сроком до трех лет. Таким образом, доходность ГЦБ МФРК стала также влиять на формирование кривой доходности на краткосрочном конце.

Рынок ГЦБ местных исполнительных органов Казахстана имеет низкую активность и незначительный объем размещения в общем объеме ГЦБ (7% – на конец 2020 года), не оказывает значительного влияния на рынок ГЦБ и на формирование кривой доходности ГЦБ.

Таким образом, ключевую роль в построении безрисковой кривой доходности со сроком до года играют ноты НБ РК, а со сроком свыше года – ГЦБ МФ РК, выпущенные в тенге. Однако низкая ликвидность этих рынков остается главным препятствием в создании адекватной кривой. Далее мы подробно рассмотрим эти рынки и причины, ограничивающие их ликвидность.

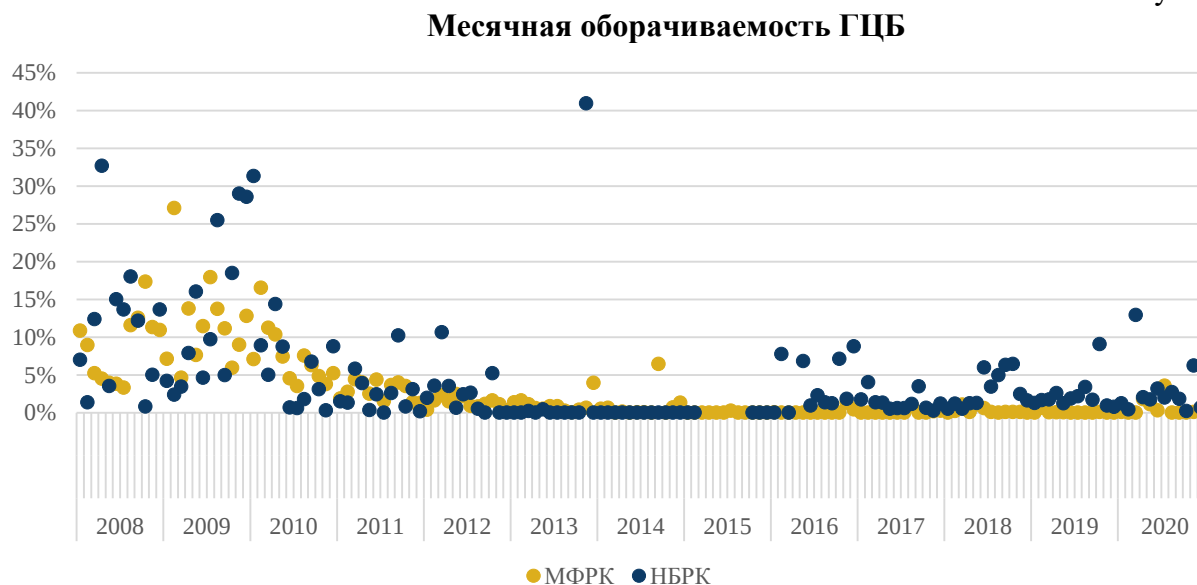
Рисунок 7



Источник: ЦДЦБ

Рынок ГЦБ РК имеет низкую ликвидность. Несмотря на значительный объем первичного рынка, сделки на вторичном рынке совершаются крайне редко. Так, максимальное значение месячного оборота ГЦБ МФ РК в 2020 году составило 3,6%, при этом 5 месяцев из 12 этот показатель был близок к нулю (рисунок 8). Ликвидность рынка ГЦБ начала снижаться в 2012 году и еще больше сократилась в 2013 году. В 2012 году наблюдалось снижение спроса на ГЦБ в связи с возросшими девальвационными ожиданиями и ослаблением регуляторных требований в части инвестирования средств пенсионных фондов [25]. В 2013 году было объявлено о консолидации пенсионных фондов и введены жесткие ограничения на инвестиционную деятельность пенсионных фондов, продажи ГЦБ из пенсионных портфелей были запрещены [19], чем и объясняется снижение ликвидности ГЦБ в 2013 году.

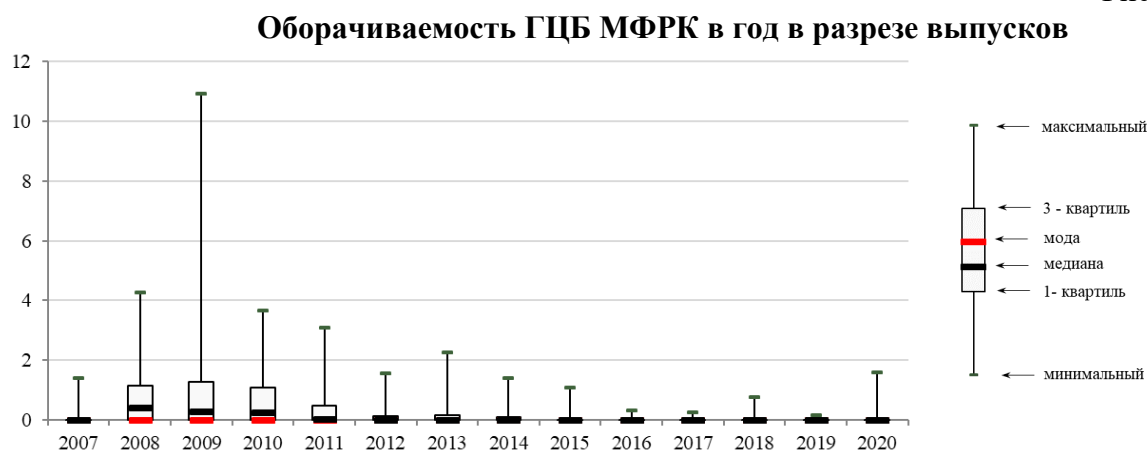
Рисунок 8



Источник: ЦДЦБ

Примечание: показатель рассчитан как отношение операций на вторичном рынке за месяц к размещенному объему на начало месяца.

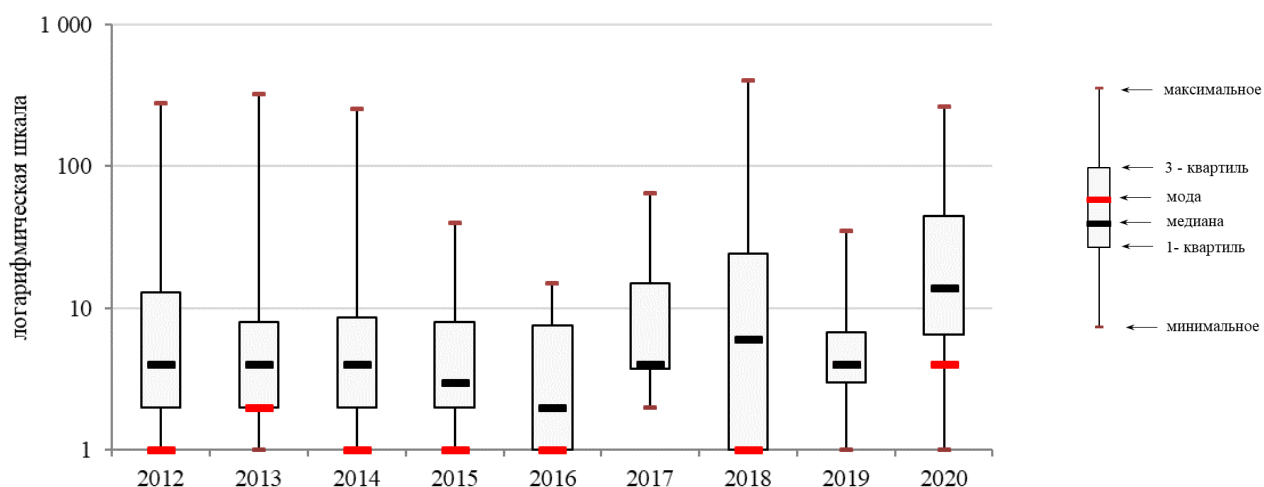
Рисунок 9



Источник: ЦДЦБ

При этом анализ количества сделок в разрезе отдельных выпусков демонстрирует положительную динамику в последние несколько лет. Увеличилось количество сделок по отдельным выпускам, что объясняется усилиями МФ РК по укрупнению объема и повышению ликвидности по отдельным выпускам с целью включения ГЦБ МФ РК в международные глобальные индексы облигаций. Тем не менее ликвидность большинства выпусков остается по-прежнему низкой, количество сделок для значительной части выпусков не превышает 1-4 сделок в год. Кроме того, по большинству выпусков, находящихся в обращении, сделки на вторичном рынке вообще не совершаются. Так, из 143 выпусков, находящихся в обращении в 2019 году, сделки на вторичном рынке совершались только по 18 выпускам, а в 2020-м – по 46 из 111 выпусков в обращении.

Рисунок 10

Количество сделок в год на вторичном рынке с ГЦБ МФ РК в разрезе выпусков

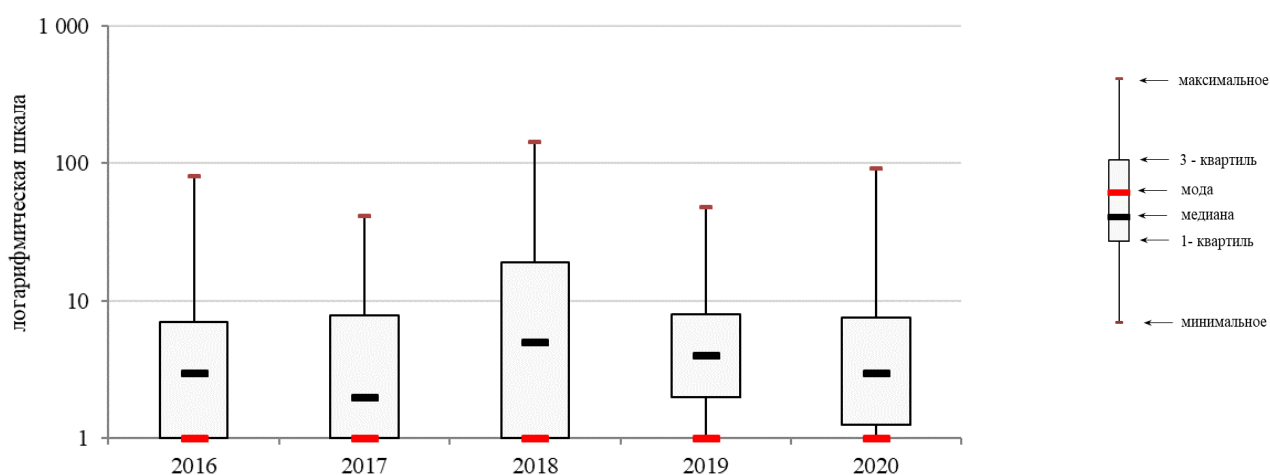
Источник: ЦДЦБ

Примечание: для построения графика учитывались только выпуски, по которым проходили сделки на вторичном рынке.

Рынок нот НБ РК

Рынок нот НБ РК имеет сравнительно большую оборачиваемость (рисунок 8), тем не менее также низколиквиден. Наиболее значительная часть выпусков нот имеет не более одной сделки в год (рисунок 11).

Рисунок 11

Количество сделок в год на вторичном рынке с нотами НБ РК в разрезе выпусков

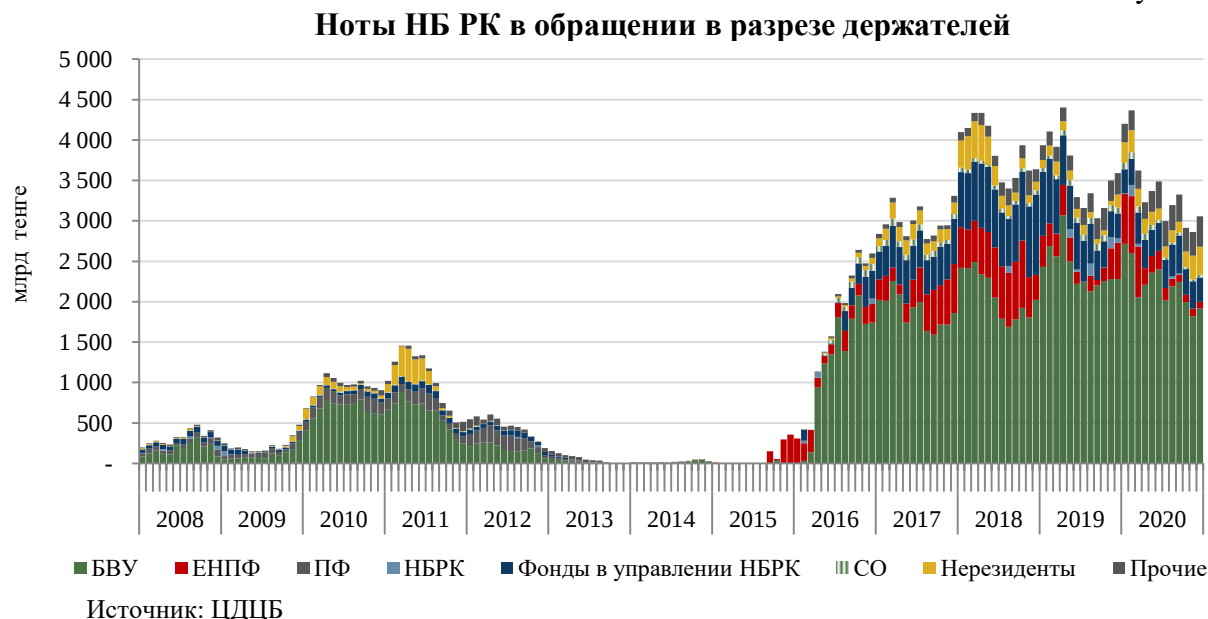
Источник: ЦДЦБ

Примечание: для построения графика учитывались только выпуски, по которым проходили сделки на вторичном рынке.

В 2013-2015 годах объемы выпущенных НБ РК нот были незначительны, и большая доля удерживалась пенсионными фондами в 2013-м и ЕНПФ в 2015 году (рисунок 12). В 2016-2017 годы в банковском секторе сложился профицит ликвидности за счет бюджетного дефицита, программ оздоровления банков и дедолларизации их обязательств. В 2015 году НБ РК перешел на режим инфляционного таргетирования и в

2016 году в рамках программы управления процентными ставками стал изымать избыточную ликвидность с помощью нот.

Рисунок 12



Доходности нот отражали ожидания добровольных участников рынка по стоимости тенговой ликвидности, а срок нот отражал короткий инвестиционный горизонт банков. Поэтому основной спрос на ноты с 2016 года формировался за счет банков второго уровня. Из добровольных инвесторов на рынке нот также участвуют нерезиденты и страховые организации, однако их доля незначительна. В результате интереса со стороны рыночных участников ноты НБ РК распределены более равномерно, большинство выпусков нот имеет большее число держателей, чем ГЦБ МФ РК, тогда как ГЦБ МФ РК имеют значительную концентрацию в разрезе держателей.

Рисунок 13



Источник: IRIS Finance

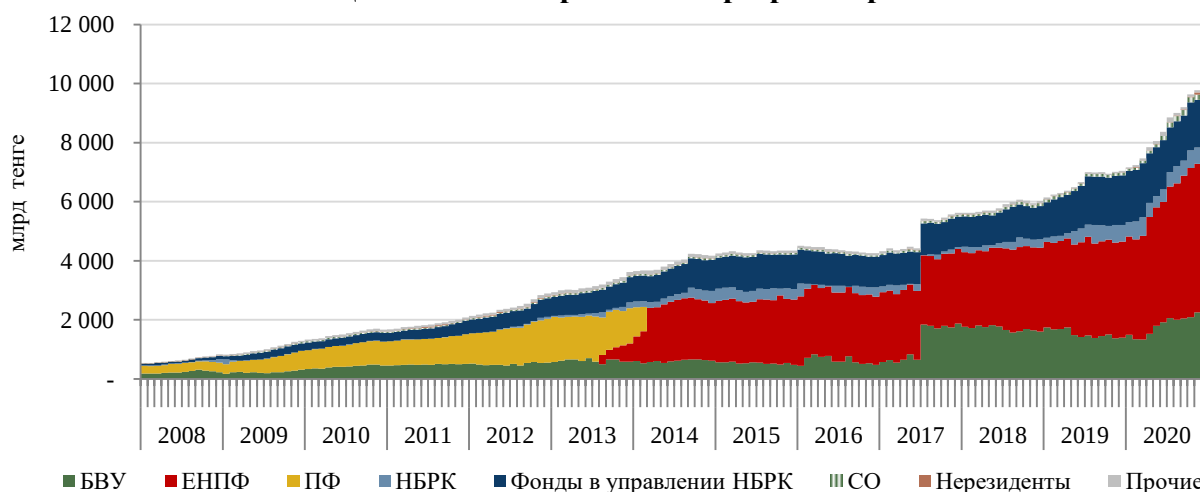
Причины низкой ликвидности. Основной причиной снижения ликвидности на рынке ГЦБ МФРК после 2014 года стало объединение пенсионных фондов в Единый накопительный пенсионный фонд (ЕНПФ). Исторически пенсионные фонды являлись

основными инвесторами на рынке ГЦБ. Так происходило в силу пруденциальных требований по достаточности капитала, согласно которым все активы пенсионных фондов взвешивались согласно их риску по версии регулятора, и в зависимости от этой величины определялся уровень достаточности капитала. ГЦБ РК считаются безрисковыми, а потому их вес равен нулю, и, соответственно, они не требуют дополнительной капитализации, чем и был обусловлен спрос на ГЦБ со стороны пенсионных фондов [31].

После объединения пенсионных фондов ЕНПФ продолжал оставаться основным инвестором на рынке ГЦБ МФ РК. Кроме того, значительная доля ГЦБ МФ РК скупалась другими фондами, находящимися в управлении НБРК. С переходом НБРК на режим инфляционного таргетирования реальная доходность ГЦБ стала положительной, что позволяло ЕНПФ формировать доходность по пенсионному портфелю выше инфляции. Несмотря на повышение реальной доходности ГЦБ, ее уровень не всегда соответствовал требованиям рыночных участников, поэтому их деятельность на рынке ГЦБ МФ РК была ограниченной. Кроме того, для рыночных участников привлекательность ГЦБ снижал риск обесценения тенге. При давлении на курс тенге доходность ГЦБ переставала отражать риск-премию по тенговым инструментам, между доходностью ГЦБ и ставками NDF обычно возникал спред [31]. С возникновением рынка нот НБРК спрос со стороны банков сместился с рынка ГЦБ на рынок нот, так как срок нот больше соответствовал инвестиционному горизонту банков, и в результате короткого срока снижались их требования по премии за ликвидность.

Рисунок 14

ГЦБ МФ РК в обращении в разрезе держателей

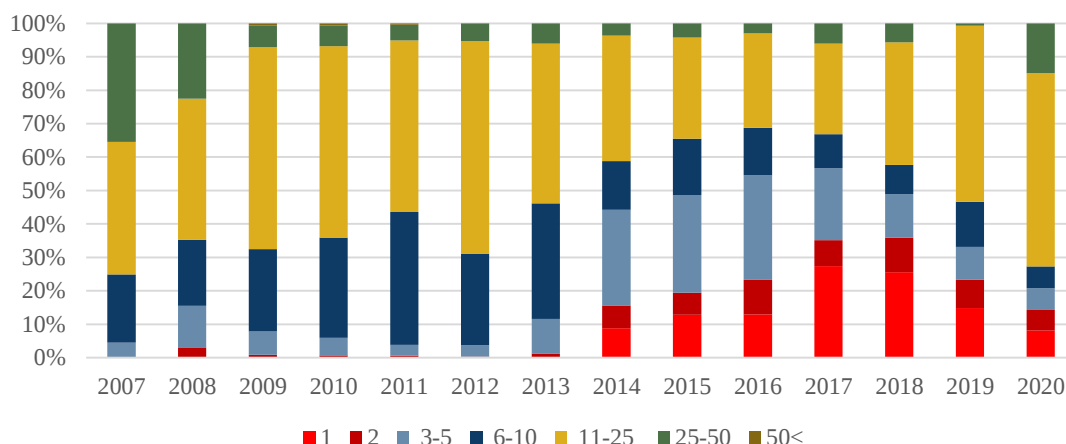


Источник: ЦДЦБ

Кроме того, одним из факторов, снижающих ликвидность рынка ГЦБ, являлась политика МФ РК по выпуску множества выпусков с мелкими объемами [27]. При таких размерах выпуска один выпуск мог оказаться в руках одного или нескольких держателей. После объединения пенсионных фондов значительная доля таких выпусков оказалась в руках ЕНПФ, что и объясняет отсутствие сделок по ним (рисунки 15, 16).

Рисунок 15

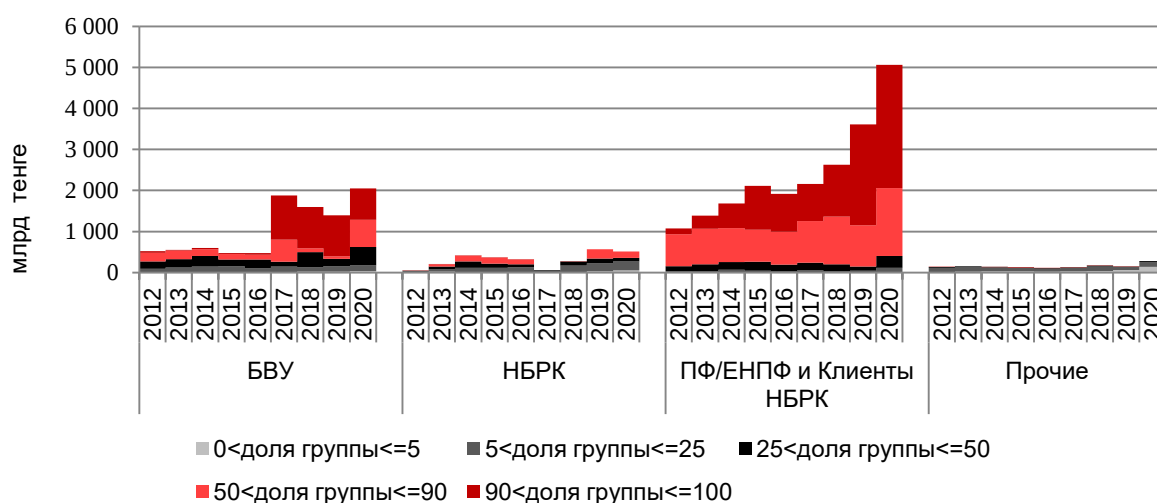
**Количество выпусков, имеющих одного держателя, возросло после
объединения пенсионных фондов**
Распределение выпусков ГЦБ МФ РК в зависимости от количества держателей



Источник: IRIS Finance

Рисунок 16

Рост концентрации ГЦБ МФРК у каптивных инвесторов
Распределение ГЦБ МФ РК, не являющихся МЕУЖКАМ, по долям видов
держателей в объеме каждого выпуска



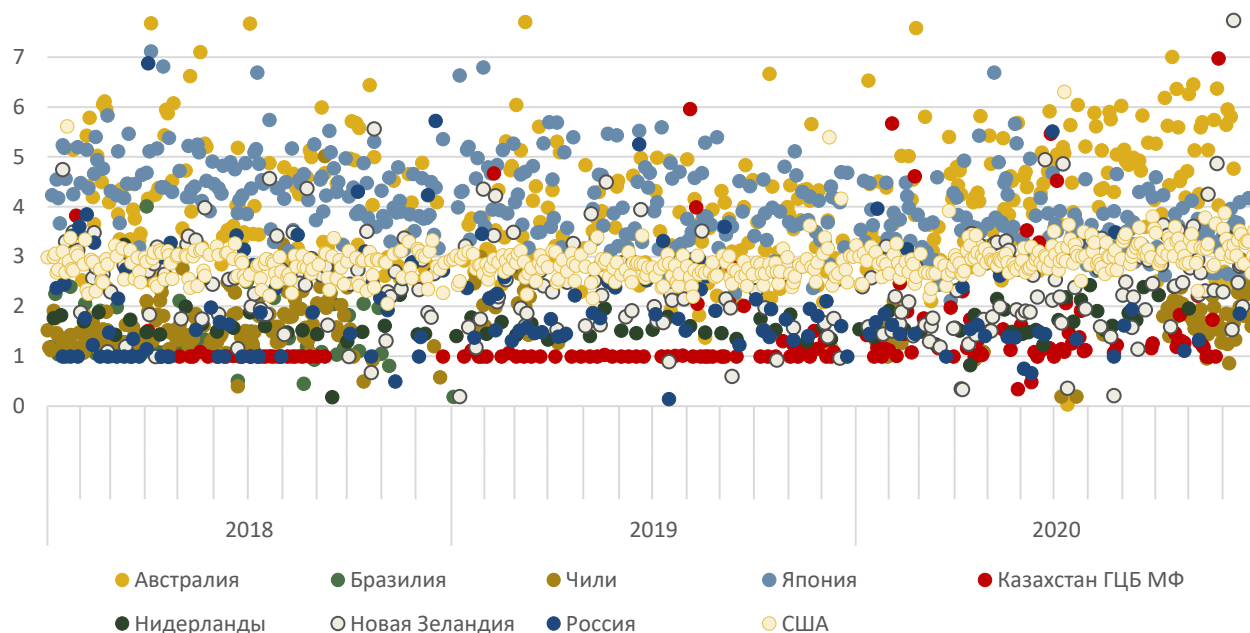
Источник: ЦДЦБ, расчеты НБ РК

Примечание: из выборки исключены ГЦБ, приобретенные системообразующим банком в июле 2017 года в рамках сделки по его оздоровлению.

Спрос на ГЦБ в странах с более развитыми рынками заметно превышает спрос на ГЦБ Казахстана (рисунок 17). При этом в большинстве случаев bid-to-cover¹ по аукционам Казахстана колеблется от 1 до 2, то есть спрос либо полностью покрывает объемы аукционов ГЦБ РК, либо превышает их, иногда почти вдвое.

¹ Bid-to-cover в данном случае отражает объем активных заявок к удовлетворенным заявкам.

Спрос к фактическому объему размещения аукционов государственных ценных бумаг



Источник: Refinitiv Eikon

Примечание: на графике показатель Bid/Cover взвешен по объему аукционов за день.

Несмотря на значительный спрос при анализе объемов заявок, анализ количества участников аукционов показывает, что спрос на ГЦБ МФ РК ограничен. В 2019 году в аукционах ГЦБ МФ РК в среднем участвовало около трех покупателей. Ограниченный спрос на государственный долг в нормальных условиях должен был привести к повышению доходности ГЦБ, однако наличие каптивных инвесторов позволяло МФ РК размещать долг ниже стоимости требуемой рыночными инвесторами. Так, основной спрос в 2019 году формировался каптивными инвесторами. Помимо неконкурентной цены, спрос со стороны рыночных участников снижался вследствие возросших сроков доплат размещения ГЦБ вследствие неравномерной структуры погашения обязательств МФ РК (рисунок 21) и их концентрации на ближайших сроках в 2018-2019 годах. МФ РК стал размещать ГЦБ со сроками доплат свыше 10 лет, чтобы распределить структуру погашения долга на более дальние сроки. Рыночные участники, в частности, банки, в этот период в основном участвовали на рынке нот НБ РК, где инвестиционный горизонт короче и больше соответствует их потребностям. Таким образом, принимая во внимание, что на вторичном рынке сделки совершаются крайне редко, а МФ РК выпускал только долгосрочные облигации, в 2018-2019 годах кривая доходности на сроках от года до 10 лет формировалась на основе сделок, которые давно потеряли актуальность. Соответственно, информационная ценность кривой доходности ГЦБ была крайне низкой, у рынка отсутствовал ориентир по стоимости безрисковых фондов на среднесрочном конце кривой доходности, что препятствовало развитию рынка капиталов в целом.

Для поддержания актуальности кривой доходности МФ РК необходимо постоянное обеспечение событий на всех сроках кривой доходности.

Спрос к фактическому объему размещения аукционов государственных ценных бумаг

Источник: КФБ

Примечание: 1) из выборки исключены ГЦБ, приобретенные системообразующим банком в июле 2017 года в рамках сделки по его оздоровлению; 2) I – период частных накопительных пенсионных фондов; II – период объединения; III – период ЕНПФ.

Напомним, что одним из факторов, снижающих ликвидность, была высокая фрагментация выпусков ГЦБ. Так, например, в 2017 году в Казахстане было 147 выпусков ГЦБ МФ РК в обращении, в то же время в Австралии количество активных выпусков составляло 32, в России – 43, а в Новой Зеландии – 13. Притом что количество участников на этих рынках гораздо больше, чем в Казахстане, а, соответственно, и глубина этих рынков значительно превышает глубину казахстанского рынка ГЦБ. По нашему мнению, снижение уровня фрагментации выпусков ГЦБ должно повысить их ликвидность. Теоретически при повышении объема и снижении их количества у одного выпуска ГЦБ будет больше держателей. Для снижения фрагментации выпусков ценных бумаг в соответствии с международной практикой необходимо проводить доразмещения выпусков «новых» (on-the-run²) выпусков ГЦБ с целью повышения их объема и выкупать «старые» (off-the-run³) выпуски ценных бумаг на ключевых точках кривой доходности. Таким образом, рыночная ликвидность будет сконцентрирована в небольшом количестве ценных бумаг, что должно повысить их ликвидность. Кроме того, проведение аукционов по действующим ГЦБ будет давать актуальную информацию по стоимости, что повысит информационную ценность кривой доходности и будет служить ориентиром по стоимости тенговых фондов для смежных рынков капитала.

В 2019 году между Правительством РК и НБ РК было заключено Соглашение о координации мер макроэкономической политики (далее – соглашение), направленное на повышение доверия рыночных участников к политике управления долгом и увеличение ликвидности рынка ГЦБ. По соглашению с Министерством финансов оно взяло на себя

² On-the-run – выпуски, размещенные наиболее поздно среди выпусков с определенным сроком доплат.

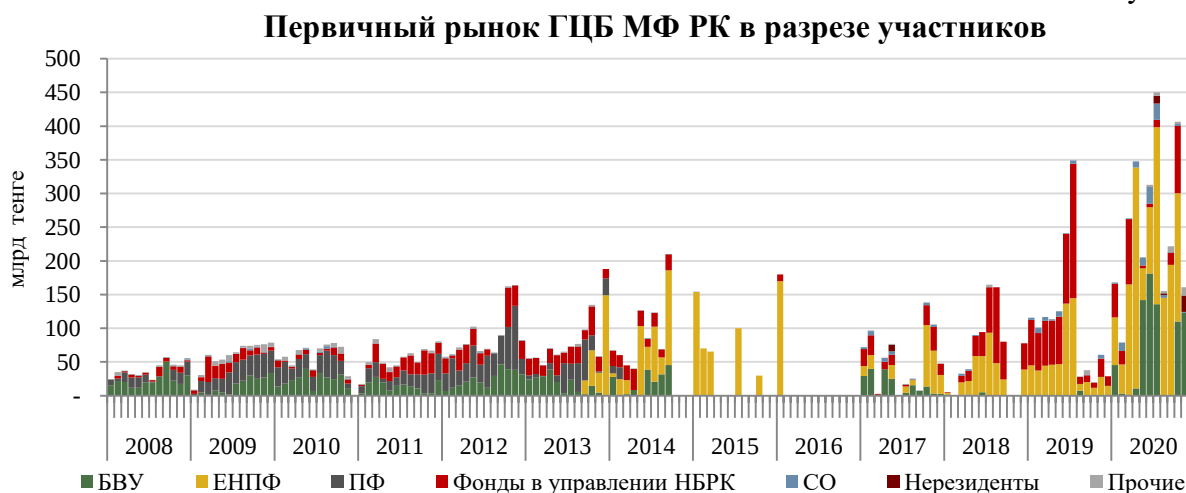
³ Off-the-run – выпуски, размещенные ранее, до наиболее позднего («нового») выпуска, и все еще находящиеся в обращении.

обязательства по публикации графиков выпусков ГЦБ на предстоящий год, а также проведению повторных аукционов по размещению ГЦБ, укрупнению выпусков ГЦБ и сокращению их количества до 30-40 выпусков в обращении в средне- и долгосрочной перспективе.

Кроме того, в середине 2019 года МФ РК изменил график выпуска ГЦБ таким образом, что начиная с августа 2019 года МФ РК стал размещать ГЦБ с разными сроками до 10 лет [27]. В результате МФ РК стал проводить дополнительные размещения ГЦБ со сроком доплатнения менее 3 лет, а в 2020 году возобновил выпуск МЕККАМ – дисконтных ГЦБ со сроком обращения до 1 года.

Предпринятые меры повысили ликвидность рынка ГЦБ в 2020 году, рыночные участники вновь стали функционировать на рынке ГЦБ МФРК, возросло количество участников аукционов. Повышение интереса рыночных участников к рынку ГЦБ МФ РК было обусловлено также тем, что в середине 2020 года НБ РК прекратил выпуск годовых нот. Таким образом, спрос, присутствовавший на рынке годовых нот НБ РК, был смещен на рынок ГЦБ МФ РК.

Рисунок 19



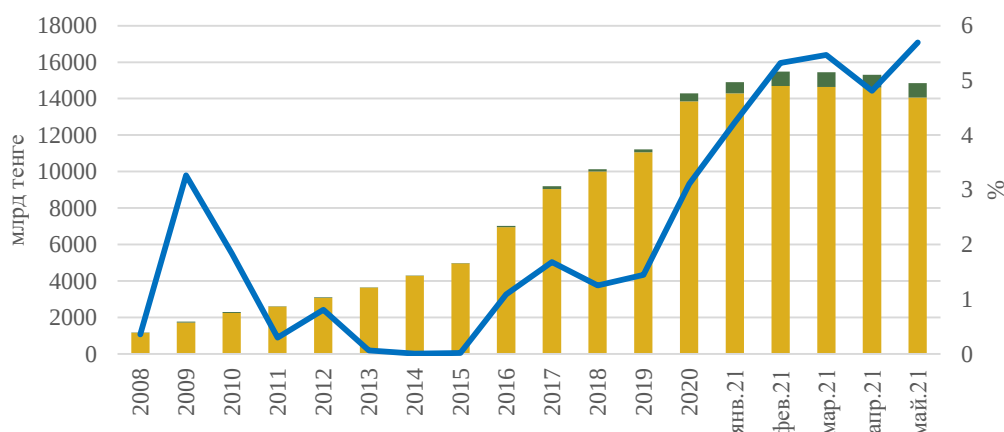
Источник: ЦДЦБ

Примечание: из выборки исключены ГЦБ, приобретенные системообразующим банком в июле 2017 года в рамках сделки по его оздоровлению.

Стоит отметить, что спрос на казахстанские ГЦБ, выпущенные в тенге, формируется в основном за счет внутреннего рынка, тогда как на более развитых рынках значительная часть спроса формируется иностранными инвесторами. Большое количество участников на развитых рынках на ГЦБ создает более конкурентную среду и позволяет повысить эффективность ценообразования.

В 2018 году НБ РК и МФ РК была начата совместная работа по повышению привлекательности рынка ГЦБ для иностранных инвесторов: 1) были начаты работы по включению казахстанских ГЦБ в индекс государственных облигаций развивающихся стран JP Morgan (GBI-EM); 2) в августе 2020 года усовершенствована международная линия расчетов до типа Delivery versus Payment по нотам Национального Банка РК и ГЦБ Министерства финансов РК. Доля нерезидентов в структуре держателей ГЦБ РК по-прежнему остается низкой, однако в результате предпринятых мер она начала увеличиваться. Наибольший рост наблюдался в 2020-2021 годы: на 1 июня 2021 доля нерезидентов среди держателей ГЦБ РК достигла 5,7%.

Рисунок 20

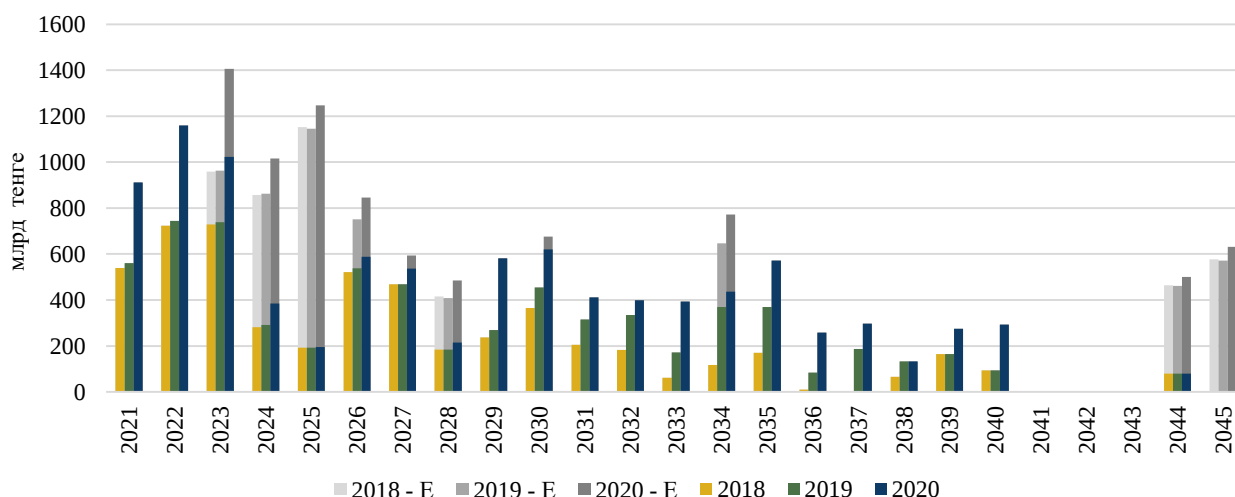
Доля нерезидентов в структуре ГЦБ РК (исключая еврооблигации)

Источник: ЦДЦБ, расчеты НБ РК

Таким образом, в прошлые годы политика управления госдолгом Казахстана в большей степени была ориентирована на текущие проблемы финансирования бюджета, а проблемы ликвидности рынка ГЦБ не являлись приоритетными. При выпуске долга МФ РК руководствовалось текущей стоимостью, не принимая в расчет риски ликвидности, связанные с концентрацией обязательств на определенных сроках, и валютные риски при выпуске обязательств в иностранной валюте.

Отсутствие четкой политики управления госдолгом наглядно иллюстрирует временная структура погашения облигаций МФ РК. Погоня за низкими ставками в валюте в середине 2010-х с последующим ослаблением тенге привело к серьезному дисбалансу графика погашения основного долга. Эта информация, очевидно, уже заложена в ставках заимствования.

Рисунок 21

График погашения основного долга по ГЦБ МФ (включая еврооблигации) по состоянию на конец периода

Источник: ЦДЦБ, расчеты НБ РК

Примечание: 1) E – обязательства МФ РК по выплате еврооблигаций; 2) сумма погашения основного долга по еврооблигациям конвертирована в тенге по официальным курсам на конец периода.

Так, в 2019 году погашение основного долга составило 3,37% от нефтяных поступлений в бюджет, а в 2020 году выросло до 4,23%. Также, согласно прогнозным

значениям поступлений по проекту республиканского бюджета на 2021-2023 годы, к 2021 году долговая нагрузка на нефтяной бюджет достигнет 9,78%, а к 2022 году составит 11,05%. Несмотря на сравнительно низкий уровень государственного долга по отношению к ВВП, отношение расходов на обслуживание долга к выпущенному долгу Казахстана (текущая стоимость долга) превышает этот показатель во многих странах. Высокая стоимость обслуживания долга обусловлена более высокими процентными ставками в Казахстане по сравнению с другими странами, а также переоценкой обязательств в иностранной валюте, произошедшей вместе с обесценением тенге.

Доверие инвесторов имеет большое значение для оптимизации госзаимствований. При этом для кредитора паттерн графика погашения часто является показателем прагматизма заемщика.

В результате совместной работы НБ РК и МФ РК отмечаются позитивные изменения в политике заимствования, направленные на повышение ликвидности рынка ГЦБ и на формировании безрисковой кривой доходности. В числе перечисленных ранее: усилия МФ РК по укрупнению и сокращению числа выпусков в обращении, работа по включению выпусков в глобальные международные индексы, реализация возможности расчетов в Clearstream, а также выпуск ГЦБ на краткосрочном и среднесрочном конце кривой доходности. В дополнение стоит отметить совместную работу НБ РК и КФБ по созданию института маркетмейкеров для рынка ГЦБ, что также может оказать позитивное влияние на ликвидность ГЦБ. Наличие маркетмейкеров дает возможность выхода из позиции по инструменту, что является немаловажным фактором для инвесторов с более коротким инвестиционным горизонтом. Кроме того, одним из необходимых условий для повышения ликвидности рынка ГЦБ и повышения рыночности принимаемых решений, по нашему мнению, является развитие рынка институциональных инвесторов. Одним из решений является передача пенсионных накоплений в инвестиционное управление частным управляющим компаниям, что будет создавать более конкурентную среду на рынке ценных бумаг.

4. Политика управления государственным долгом

Относительно низкое отношение государственного долга к ВВП в Казахстане не означает снижения актуальности задач перехода к политике активного управления государственным долгом. Динамика развития рынка государственных ценных бумаг в малых открытых экономиках в целом и в Казахстане в частности складывается преимущественно из основных макроэкономических условий и предпосылок для внутренней экономики. В Казахстане, где основным каналом покрытия дефицита бюджета все еще является высокий и растущий за последние годы трансферт из Национального фонда РК, соответственно, возрастает зависимость финансирования фискального баланса от нефтяных поступлений. Это повышает процикличность фискальной политики (повышение госрасходов в период высоких цен на нефть), тем самым сужает потенциал фискального пространства в качестве буфера в кризисные периоды. Также преобладание такой тенденции косвенно сказывается на долгосрочном снижении инвестиционного рейтинга, что негативно влияет на возможность привлечения иностранного капитала во внутренний долг [30]. В свою очередь присутствие нерезидентов в качестве держателей госдолга может улучшать диверсификацию инвесторской базы, что тем самым способствует повышению ликвидности государственных ценных бумаг – центральной задачи скоординированной макроэкономической политики.

Поэтому эффективное развитие рынка государственных ценных бумаг является приоритетной задачей как фискального рукава правительства, так и ее денежно-кредитной составляющей. Результативное проведение трансмиссии процентных ставок денежно-кредитной политики опирается на корректное функционирование конкурентной суверенной кривой доходности, которое также необходимо для ценообразования рисков активов. В дополнение к этому дальнейшее развитие рынка ГЦБ должно быть

направлено на его развитие как первичного звена трансмиссионного механизма денежно-кредитной политики: на повышение эффективности трансмиссии ставок денежно-кредитной политики на ставки ГЦБ и непосредственное влияние на условия фондирования в экономике [16]. Кроме того, исследования в области управления госдолгом в развивающихся странах показывают, что увеличение выпуска долга в национальной валюте способствует большей ответной реакции рынка на изменения в денежно-кредитной политике.

Планомерное и всестороннее развитие рынка государственного долга, помимо своей первостепенной макроэкономической задачи, должно включать в себя также и задачи структуры расходной части бюджета и целей, под которые происходит выпуск госдолга. Одним из перспективных направлений развития активного управления госдолгом может служить оценка инвестиционной направленности займов – секторальный анализ средств, которые финансируются через выпуск государственных ценных бумаг [3]. Таким образом, реализация программ, обеспечивающих уклон в сторону наукоемкого и капиталоемкого производства, способствует более качественному и высокому темпу экономического роста через увеличение производительности труда, что в свою очередь волнообразно повышает спрос на дальнейшие объемы государственных заимствований и их эффективное функционирование.

По нашему мнению, проблематика целей фискального ведомства заключается в приоритете краткосрочных целей над целями реального предназначения: долгосрочное стабильное улучшение благосостояния страны, снижение заимствования из Национального фонда – фонда резервов для будущих поколений. Таким образом, с точки зрения правительства целью эффективности управления госдолгом должна являться не текущая стоимость заимствования, а логически обоснованная долгосрочная структура долга и снижение, по сути, бремени для будущих поколений. В этом контексте необходимо нивелировать существующий конфликт интересов: для правительства рынок должен быть надежным, прозрачным и стабильным каналом финансирования, а для финансового регулятора – эффективной площадкой реализации денежно-кредитной политики.

Действующая временная структура погашения ГЦБ, на текущий момент подверженная значительным рискам ликвидности и валютным рискам, может быть оптимизирована одним из основных методов обратной покупки:

1. Обратный выкуп. Аукцион проводится так же, как и обычный аукцион, за исключением того, что участники торгов подают заявки на продажу аналогично механизму первичного рынка.

2. Прямая покупка. Этот метод выкупа долга может осуществляться двумя способами. Первый – это выкуп ГЦБ на вторичном биржевом рынке. Второй – это покупка на внебиржевом рынке.

3. Замена или рефинансирование, когда одна ценная бумага предоставляется в качестве оплаты за покупку другой ценной бумаги. Во время каждой сделки по обратному выкупу важно учитывать доходность отсечения выкупаемой ГЦБ. Оценка кривой доходности на вторичном рынке даст представление о справедливой рыночной стоимости бумаги. В этом контексте теоретическая модель кривой доходности должна стать лучшей метрикой при определении реальных цен отсечения.

Консолидация выпусков ГЦБ, а также поддержание крупных выпусков повышают ликвидность рынка. Неэффективные бумаги с точки зрения конкурентного ценообразования повышают недоверие со стороны рыночных участников, что в целом мешает развитию рынка до его потенциала и препятствует росту капитализации других площадок обмена капитала. Наоборот, поддержание крупных «эффективных» бумаг будет поддерживать и повышать ликвидность рынка ГЦБ. Это повышает эффективность системы фондирования, как с точки зрения долгосрочных целей для правительства, так и

менее долгосрочных целей монетарных властей, что возможно только при их тесной координации в части активного управления госдолгом.

5. Заключение

В целом ликвидность на рынке ценных бумаг в последние два года постепенно улучшается. В этот период важным фактором стало изменение практики эмиссии долга. Это включало в себя обязательство по созданию крупных контрольных выпусков, которые помогли росту ликвидности. Рынок ГЦБ также поддерживался изменениями в структуре участников рынка, что усиливает дилерскую конкуренцию. Переход на новую модель построения кривой доходности, повышающий уровень прозрачности, также предлагает потенциал повышения ликвидности в будущем. Степень ликвидности рынка продолжает существенно зависеть от локальной конъюнктуры: в периоды нестабильности процентных ставок ликвидность стремительно падает.

С точки зрения потенциала развития рынок РК, кажется, выгодно отличается от выборки сравниваемых стран. Тем не менее с точки зрения структурных факторов, которые создают обстановку, степень фрагментации долговых обязательств, паттерн графика погашения и относительная неразвитость фондового рынка могут препятствовать дальнейшему повышению ликвидности.

Необходимость рациональной политики управления госдолгом обуславливает актуальность исследований международных практик и разработок моделей в этом направлении. Переход на новую модель политики управления госдолгом обеспечит ее устойчивость и преемственность. В этой статье дается краткий обзор структуры ликвидности рынка ГЦБ, помогающий понять и организовать это многогранное решение.

Список использованной литературы

1. Broto C., Lamas M., «Measuring market liquidity in US fixed income markets: A new synthetic indicator» // Working Paper No. 1608, Bank of Spain, 2016.
2. Elliot D., «Market liquidity: A primer. The Brookings Institution» // Working paper, The Brookings Institution 2015.
3. Fiscal policy, public debt and monetary policy in emerging market economies, BIS Papers, №67, 2012. <https://www.bis.org/publ/bppdf/bispap67.pdf>.
4. Gehde-Trapp M., Schuster Ph., Uhrig-Homburg M., «The Term Structure of Bond Liquidity» // Journal of Financial and Quantitative Analysis, vol. 53, No. 5, p. 2161-2197, 2018.
5. Global Financial Stability Report: Markets in the Time of COVID-19. International Monetary Fund. Washington, DC., April 2020. <https://www.imf.org/en/Publications/GFSR/Issues/2020/04/14/global-financial-stability-report-april-2020>.
6. Grobety M., «Government debt and growth: The role of liquidity» // Journal of International Money and Finance, vol. 83, p 1-22, May, 2018.
7. Gulliver J., «Money Market Funds and the 2020 COVID Crisis» // Committee on Capital Markets Regulation, May 2021. <https://www.capmksreg.org/wp-content/uploads/2021/05/CCMR-Prime-MMFs-and-COVID-05.24.2021-.pdf>.
8. Kazakhstan: Reversing Productivity Stagnation – Country Economic Memorandum // [World Bank Document](#).
9. Le Saout, E., «Intégration du risque de liquidité dans les modèles de valeur en risque» // Banque et Marchés №61, p. 15-25, novembre 2002.
10. Muranaga J., T. Shimizu, «Market Microstructure and Market Liquidity» // Working paper of Institute for Monetary Economic Studies, Bank of Japan, 1997.
11. Nazlioglu S., Gupta R., Bouri E. «Movements in International Bond Markets: The Role of Oil Prices» // Working paper № 2019:35, International Review of Economics & Finance. – 2020.

12. Phillip R.D. Anderson, A. Velandia-Rubiano, A.C. Silva, «Public Debt Management in Emerging Market Economies: Has this Time Been Different?» // World Bank Policy Research Working Paper, No. 5399, April, 2016. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1658774.
13. V. Premathiratne, «The Emerging Asia Pacific Capital Markets Sri Lanka» // 2021 CFA Institute Research Foundation, May, 2021.
14. Semernina Y., Yakunina A., Yakunin S., «Some trends of the Russian sub-federal bond market development in the presence of the global challenges» // Advances in Economics, Business and Management Research (AEBMR), vol. 39, p. 162–165, 2018.
15. Yakunina A.V., Semernina Y.V., Yakunin S.V., Ermakova E., «Impact of foreign economic sanctions on the Russian market of bonds denominated in foreign currencies» // SHS Web Conf. Innovative Economic Symposium Milestones and Trends of World Economy (IES 2018), vol. 61, 2019.
16. Алексеева И.А., Петрова И.А., «Эволюция и перспективы развития рынка облигаций федеральных займов» // Известия Байкальского государственного университета, Т. 28, № 4, 2018. <https://cyberleninka.ru/article/n/evolyutsiya-i-perspektivy-razvitiya-rynka-obligatsiy-federalnyh-zaymov>.
17. Алехин Б.И., «Размещение государственных облигаций сквозь призму микроструктуры рынка» // Финансовый журнал, № 6, стр. 95–108, 2018.
18. Белокопытов А.В., Миронкина А.Ю., «Облигации федерального займа для населения» // Экономический журнал, стр. 54–62, 2017. <https://cyberleninka.ru/article/n/obligatsii-federalnogo-zayma-dlya-naseleniya>.
19. Мустафин Е.Т., «Долговой рынок в период реформы НПФ» // Halyk Finance, январь, 2013. <http://history.halykfinance.kz/ru/site/index/research/report:81987>.
20. Мустафин Е.Т., «Кривая доходности Казахстана: Падение спроса на госдолг в ожидании пруденциальных послаблений» // Halyk Finance, сентябрь, 2012. <http://history.halykfinance.kz/ru/site/index/research/report:81322>.
21. Мустафин Е.Т., Хакимжанов С.Т., «Риск ликвидности, обменный курс и стоимость фондирования в тенге» // Halyk Finance, январь, 2014. <http://history.halykfinance.kz/ru/site/index/research/report:87302>.
22. Огурцова Е.В., Комков И.В., «Развитие внутреннего рынка государственного долга России: факторы формирования, перспективы изменения» // Известия Саратовского университета, серия «Экономика. Управление. Право», Т. 19, вып. 3, 2019. <https://doi.org/10.18500/1994-2540-2019-19-3-246-256>.
23. Оразалин Р., «Трансмиссионный механизм денежно-кредитной политики в Республике Казахстан» // Экономическое исследование №2018-3, Департамент исследований и статистики, Национальный Банк Республики Казахстан, 2018.
24. Основные направления развития финансового рынка Российской Федерации на период 2019–2021 годов // Банк России. https://cbr.ru/content/document/file/71220/main_directions.pdf.
25. Отчет о финансовой стабильности Казахстана 2012. Национальный банк Республики Казахстан. Алматы, декабрь, 2012. <https://nationalbank.kz/file/download/26119>.
26. Отчет о финансовой стабильности Казахстана 2015 – 2017. Национальный банк Республики Казахстан. Алматы, 2018. <https://www.nationalbank.kz/ru/news/otchet-o-finansovoy-stabilnosti/rubrics/235>.
27. Отчет о финансовой стабильности Казахстана 2018 – 1П 2019. Национальный банк Республики Казахстан. Алматы, 2020. <https://nationalbank.kz/ru/news/otchet-o-finansovoy-stabilnosti/rubrics/236>.
28. Покровская Н.Н., Куриленко М.В., «Рынок государственных ценных бумаг в России: проблемы и перспективы развития» // Интернет-журнал «Науковедение», том 8, № 4, 2016. <http://naukovedenie.ru/PDF/34EVN416.pdf>.

29. Рыжановская Л.Ю., Чистова В.Е., «Проблемы ликвидности рынка государственных ценных бумаг» // «Финансовый журнал», №3, стр. 19–32, 2011.

30. Стахович Л.В., Шахназарян Г.Э., «Проблемы ликвидности современного российского рынка государственных ценных бумаг» // «Финансовая аналитика: проблемы и решения», № 1 (1), стр. 34–43, 2008.

31. Хакимжанов С., Мустафин Е., Кубенбаев О., Атабек Д., «Построение кривой доходности на рынке с низкой ликвидностью» // журнал «Деньги и кредит», №4, декабрь, 2019.

ПЕРСПЕКТИВЫ И РАЗВИТИЕ НЕБАНКОВСКИХ ФИНАНСОВЫХ ОРГАНИЗАЦИЙ

Бакиров А.М. – начальник управления развития банков и небанковских организаций департамента развития финансовых организаций Национального Банка Республики Казахстан.

Бижанов Б.А. – заместитель начальника управления развития банков и небанковских организаций департамента развития финансовых организаций Национального Банка Республики Казахстан.

Завалина А.В. – главный специалист управления развития банков и небанковских организаций департамента развития финансовых организаций Национального Банка Республики Казахстан.

Садыбаева Д.С. – главный специалист управления развития банков и небанковских организаций департамента развития финансовых организаций Национального Банка Республики Казахстан.

Статья посвящена роли и значению функционирования небанковских организаций в Республике Казахстан, в частности, микрофинансовых организаций, и описывает возможные направления их развития и перспективы деятельности.

Ключевые слова: небанковские организации, микрофинансовые организации, деятельность микрофинансовых организаций в Республике Казахстан.

JEL-классификация: G21, G23.

1. Характеристика кредитного рынка Казахстана

Вопросы, связанные с формированием и развитием финансовой, в том числе кредитной, системы, всегда имели большое значение и актуальность. Небанковские организации как участники финансового рынка осуществляют важнейшую роль в развитии финансовой системы. Внедрение новых механизмов оптимизации деятельности небанковских организаций способствует повышению эффективности рынка кредитных услуг, созданию субъектов финансового рынка, отвечающих современным мировым стандартам и способных максимально удовлетворять потребности рынка по оказанию финансовых услуг.

В связи с этим изучение текущего положения и принятие обоснованных решений в перспективе развития финансовой системы страны являются необходимыми условиями для обеспечения эффективной и продуктивной работы финансовых, в том числе кредитных институтов.

Таким образом, рассмотрение особенностей и перспектив развития небанковских организаций Республики Казахстан является достаточно актуальной темой исследования.

В целом на сегодняшний день банковский сектор в Республике Казахстан играет доминирующую роль в финансовой системе по оказываемым услугам, в том числе по выдаче займов.

Каждый год динамика объема займов растет: только за 2020 год банками выдано кредитов на сумму 1 трлн тенге.

Рост банковских кредитов в 2020 году был обусловлен в частности антикризисными мерами государства в рамках кредитования экономики и субсидирования займов. Основными заемщиками банков на рынке кредитных услуг являются физические лица. На их долю по состоянию на 1 января 2021 года приходилось 45,1% всех кредитов, с

начала года данный показатель вырос на 796 млрд тенге. При этом кредиты юридическим лицам снизились на 267 млрд тенге, или на 24,3%. Это свидетельствует о том, что в последнее время акцент сместился на розничное кредитование физических лиц, платежеспособность которых растет.

В основном потребительским кредитованием населения занимаются крупные банки с большой клиентской базой, сетью филиалов и кассовыми отделениями по стране.

В перспективе развития конкуренции на рынке финансовых услуг по обеспечению доступного кредитования не только в крупных городах, но и в регионах страны, приоритетным является развитие таких небанковских кредитных организаций, как микрофинансовые организации, для обслуживания широкого круга потребителей большим спектром предоставляемых кредитных продуктов.

Таким образом, основное отличие небанковских организаций от коммерческих банков заключается:

- в спектре оказываемых банковских услуг;
- в регулировании их деятельности;
- в возможности открытия филиалов и дочерних организаций и др.

2. Микрофинансовые организации в РК

По статистике, 67% заемщиков микрофинансовых организаций проживают в сельской местности и не интересны банкам в силу географического местоположения, отсутствия ликвидных залогов, невысокой финансовой и цифровой грамотности.

Развитие рынка микрофинансовых организаций способствует развитию агропромышленности, в том числе таких сфер, как животноводство, пчеловодство, и т.д., являющихся малопривлекательными для банков второго уровня. Основной целевой аудиторией микрокредитования являются сельская местность и малый бизнес. Микрофинансовые организации на протяжении более 20 лет работают с такой категорией населения и знают их потребности, обладают собственной методикой микрокредитования, основанной на лучших мировых практиках, материально-технической базой и программным обеспечением, обученным и квалифицированным персоналом, широкой сетью отделений и офисов в регионах. Микрофинансовые организации поддерживают предпринимателей-новичков, их стартап-проекты, предоставляя стартовый капитал тем, кто не имеет доступа к банковским кредитам. Так микрофинансовые организации расширяют возможности и создают благоприятные условия для них.

Таким образом, формирование сети микрофинансовых организаций в регионах страны является основной предпосылкой расширения доступа к финансовым ресурсам для безработного, самозанятого населения, субъектов малого и среднего бизнеса.

По состоянию на 1 апреля 2021 года микрофинансовый сектор Республики Казахстан представлен 179 организациями, осуществляющими микрофинансовую деятельность (получившие лицензию). Активы микрофинансовых организаций составили 574,3 млрд тенге, с положительной динамикой на 57 млрд тенге, общий ссудный портфель микрофинансовых организаций составил 463,8 млрд тенге, увеличившись с начала года на 48,2 млрд тенге.

Исторически микрофинансовые организации развивались эволюционно – от нерегулируемого сектора до деятельности, подлежащей лицензированию.

В целях формирования прозрачного рынка микрофинансовых организаций и увеличения к ним доверия, повышения качества и доступности предоставляемых ими услуг в Казахстане сформирована нормативная правовая база, которая совершенствуется в достаточной мере.

В целях защиты прав потребителей финансовых услуг Законом Республики Казахстан от 3 июля 2020 года «О внесении изменений и дополнений в некоторые законодательные акты Республики Казахстан по вопросам ипотечных займов в

иностранной валюте, совершенствования регулирования субъектов» с 1 января 2021 года в Республике Казахстан введено лицензирование микрофинансовой деятельности.

Лицензирование микрофинансовой деятельности позволило признать субъектов микрофинансовой деятельности финансовыми организациями и создало возможности для единого законодательного определения финансовых услуг и финансового рынка, установило более высокие требования к их деятельности, к их финансовой устойчивости, в целях поддержания добросовестной конкуренции на финансовом рынке создало равноправные условия для их деятельности, что способствует защите прав и интересов потребителей финансовых услуг.

При этом для защиты прав потребителей финансовых услуг необходимо, чтобы финансовая доступность услуг и финансовая грамотность граждан развивались параллельно.

Целями развития деятельности микрофинансовых организаций являются обеспечение доступными заемными ресурсами населения, среднего и малого бизнеса, индивидуального предпринимательства, а также повышение уровня конкурентоспособности и финансовой устойчивости данных организаций.

Растущий уровень конкуренции, высокие требования к качеству предоставляемых услуг обуславливают постоянное стремление кредитных организаций к достижению высоких стандартов качества, завоеванию лидирующего положения в финансовой сфере, получению конкурентных преимуществ.

Таким образом, микрофинансовые организации должны поспособствовать заполнению пробелов на рынке финансовых услуг, возможно, образуемых банками второго уровня.

При этом одной из ключевых проблем в микрокредитовании являются высокие ставки вознаграждения. На сегодняшний день среднее значение годовой эффективной ставки вознаграждения по микрокредитам по рынку микрофинансовых организаций составляет около 38%.

Высокие ставки вознаграждения обусловлены, во-первых, ставкой фондирования в размере 17%. Согласно сведениям ОЮЛ «Ассоциация микрофинансовых организаций Казахстана», микрофинансовые организации испытывают проблему привлечения внутри Казахстана ресурсов, используемых для своей основной деятельности. Отечественные банки требуют от микрофинансовых организаций в обеспечение ликвидные залогов, которые у них отсутствуют или имеются в недостаточном количестве. В связи с этим микрофинансовые организации вынуждены обращаться к иностранным кредиторам, предоставляющим в основном беззалоговые кредиты. Около 70% всех микрокредитов, предоставляемых в Казахстане, выдаются за счет зарубежного фондирования. Во-вторых, необходимо учитывать операционные расходы, связанные с выдачей и обслуживанием микрокредитов, что составляет около 14%. Основное преимущество микрофинансовых организаций заключается в осуществлении деятельности в отдаленных регионах Казахстана, где слабо представлены банки. Это подразумевает под собой непосредственный контакт с заемщиками, оценку платежеспособности клиента путем посещения места его бизнеса, систематический мониторинг микрокредита. Как указывалось выше, по статистике, 67% заемщиков микрофинансовых организаций являются жителями сельской местности. Содержать разветвленную сеть отделений и большое количество обученного персонала достаточно дорого. Поэтому в операционных расходах основную долю занимают расходы на персонал (фонд оплаты труда, налоги с заработной платы), амортизацию основных средств и нематериальных активов. Дополнительно формируются резервы по кредитным рискам, на долю которых приходится 18% всех расходов. Прочие расходы (аренда и охрана помещений, ремонт и содержание офисов, коммунальные платежи, связь и интернет, обучение персонала, маркетинговые расходы, содержание автотранспорта и пр.) составляют 27%.

Таким образом, все перечисленные компоненты в общей сложности формируют стоимость ставки в 34%. Если средняя ставка вознаграждения складывается в размере 38%, то очевидно, что прибыль формируется в размере 4%.

3. Предложения по развитию микрофинансового рынка

Компенсация операционных расходов микрофинансовых организаций могла бы полностью обеспечить потребность населения в сельских регионах в финансировании с низкими процентными ставками. В связи с этим для действующих микрофинансовых организаций предлагается предусмотреть стимулы для снижения издержек в микрокредитовании, что позволит снизить ставки для конечных потребителей. Одним из возможных стимулов могут стать налоговые льготы для частных организаций, предоставляющих микрокредиты начинающим предпринимателям, сельхозтоваропроизводителям и другим субъектам малого и среднего бизнеса в регионах Казахстана.

Сокращение операционных расходов возможно за счет внедрения инновационных технологий в свою деятельность, расширения дистанционных каналов обслуживания клиентов, автоматизации кредитных процессов. При этом для таких изменений необходимы ресурсы, время и благоприятное правовое регулирование.

Реализация указанных мер позволила бы снизить расходы микрофинансовых организаций, а также полностью обеспечить потребность населения в финансировании с низкими процентными ставками.

В настоящее время отмечается низкая заинтересованность банков в финансировании субъектов микро- и малого бизнеса в рамках программы субсидирования ставки вознаграждения по кредитам. В связи с этим также предлагается привлечь к программам по субсидированию процентной ставки микрофинансовые организации.

В соответствии с Национальным планом развития Республики Казахстан до 2025 года для развития финансовой инклюзии и доступности базовых финансовых услуг для населения, микро- и малого бизнеса необходимо активизировать сектор небанковского кредитования и микрофинансирования.

В целях увеличения доступности финансовых услуг для населения, противодействия проведению незаконных финансовых и мошеннических операций и дальнейшего сокращения теневой экономики также необходимо улучшить систему регулирования и надзора сектора небанковских микрофинансовых организаций.

Для повышения эффективности реализации государственных мер поддержки и максимального охвата ими большей части субъектов микро- и малого бизнеса предлагается вовлечение небанковского сектора в реализацию антикризисных программ поддержки приоритетных отраслей экономики, прежде всего, агропромышленного комплекса.

Как известно, в результате реализации антикризисных мер в экономике у банков возникла избыточная денежная ликвидность. При этом существуют ниши, в которые эти средства не поступают. Банки второго уровня не заинтересованы во вложениях в небольшие проекты в регионах Казахстана.

В соответствии с Посланием Главы государства К.Токаева от 1 сентября 2021 года «Единство народа и системные реформы – прочная основа процветания страны» необходимо задействовать потенциал микрофинансовых организаций, работающих на местах, знающих своих клиентов, их бизнес и возможности.

С учетом изложенного стоит отметить, что инфраструктура кредитных отношений играет важную роль в функционировании финансовых рынков, так как в процессе осуществления своей деятельности решает макроэкономические финансовые задачи, является сегментом национальной финансовой системы.

Список использованной литературы

1. Закон Республики Казахстан от 3 июля 2020 года «О внесении изменений и дополнений в некоторые законодательные акты Республики Казахстан по вопросам ипотечных займов в иностранной валюте, совершенствования регулирования субъектов».
2. Статистические данные с сайта Национального Банка Республики Казахстан и Агентства Республики Казахстан по регулированию и развитию финансового рынка.
3. Послание Главы государства К.Токаева от 1 сентября 2021 года «Единство народа и системные реформы – прочная основа процветания страны».