



НАЦИОНАЛЬНЫЙ БАНК КАЗАХСТАНА

# СИСТЕМА СЕЛЕКТИВНО - КОМБИНИРОВАННОГО ПРОГНОЗА ИНФЛЯЦИИ (SSCIF)

Департамент денежно-кредитной политики  
Экономическое исследование №2024-13  
Рабочая статья

Адилханова З.  
Ержан И.

Экономические исследования и аналитические записки Национального Банка Республики Казахстан (далее – НБРК) предназначены для распространения результатов исследований НБРК, а также других научно-исследовательских работ сотрудников НБРК. Экономические исследования распространяются для стимулирования дискуссий.

Рабочие статьи сотрудников НБРК отражают незавершённые исследования по состоянию на дату публикации и предназначены для обсуждения, получения комментариев и замечаний. Мнения и суждения, представленные в статье, отражают мнение авторов и не должны восприниматься как отражающие взгляды НБРК или его руководства.

## **Система селективно - комбинированного прогноза инфляции (SSCIF)**

**Декабрь 2024**

**NBRK – WP – 2024 – 13**

# Система селективно - комбинированного прогноза инфляции (SSCIF)

Адилханова Зарина<sup>1</sup>  
Ержан Ислам<sup>2</sup>

## Аннотация

В условиях нестабильной макроэкономической среды повышение точности прогнозирования инфляции является приоритетной задачей для центральных банков, особенно тех, которые придерживаются режима инфляционного таргетирования. Традиционные эконометрические модели сталкиваются с ограничениями при учёте волатильности, внешних шоков и нелинейных взаимосвязей. Данное исследование направлено на улучшение прогнозирования инфляции путём интеграции методов машинного обучения в существующую систему селективно-комбинированного прогнозирования инфляции. Включение таких алгоритмов, как Ridge Regression, Lasso Regression и Elastic Net, позволяет выявлять сложные паттерны в макроэкономических данных и повышать точность прогнозов.

Сравнительный анализ прогнозов, полученных с использованием традиционных эконометрических моделей (OLS, LTAR, BVAR, RW) и алгоритмов машинного обучения, показывает, что гибридный подход значительно снижает ошибки прогнозирования и повышает надёжность прогнозов в краткосрочном периоде. Полученные результаты могут внести вклад в совершенствование инструментов макроэкономического прогнозирования и развитие более эффективной денежно-кредитной политики, поддерживая качество принятия решений центральными банками.

**Ключевые слова:** инфляция, прогнозирование, индекс потребительских цен, модель, машинное обучение, эконометрические модели, точность прогнозов.

**Классификация JEL:** E31, E37, C52, C61.

---

<sup>1</sup> Адилханова Зарина – главный специалист-аналитик Управления макроэкономических исследований и прогнозирования Департамента денежно-кредитной политики Национального Банка Республики Казахстан.  
E-mail: [zarina.adilkhanova@nationalbank.kz](mailto:zarina.adilkhanova@nationalbank.kz)

<sup>2</sup> Ержан Ислам – заместитель начальника Управления макроэкономических исследований и прогнозирования Департамента денежно-кредитной политики Национального Банка Республики Казахстан  
E-mail: [islam.yerzhan@nationalbank.kz](mailto:islam.yerzhan@nationalbank.kz)

## Содержание

<b>1</b>	<b>Введение</b> .....	<b>5</b>
<b>2</b>	<b>Обзор литературы</b> .....	<b>6</b>
<b>3</b>	<b>Описание данных и методология</b> .....	<b>9</b>
<b>3.1</b>	<b>Описание системы селективного прогнозирования инфляции и данных</b> .....	<b>9</b>
<b>3.2</b>	<b>Обновление существующих эконометрических моделей</b> .....	<b>10</b>
<b>3.3</b>	<b>Использование моделей машинного обучения</b> .....	<b>11</b>
<b>3.3.1</b>	<b>Ridge Regression</b> .....	<b>11</b>
<b>3.3.2</b>	<b>Lasso Regression</b> .....	<b>13</b>
<b>3.3.3</b>	<b>Elastic Net</b> .....	<b>14</b>
<b>4</b>	<b>Результаты</b> .....	<b>15</b>
<b>4.1</b>	<b>Улучшение эконометрических моделей</b> .....	<b>15</b>
<b>4.2</b>	<b>Использование моделей машинного обучения</b> .....	<b>19</b>
<b>5</b>	<b>Выводы</b> .....	<b>22</b>
<b>6</b>	<b>Список использованной литературы</b> .....	<b>24</b>
<b>7</b>	<b>Приложение</b> .....	<b>25</b>

## 1 Введение

В условиях глобальной экономической нестабильности и нарастающей неопределённости способность точно прогнозировать инфляцию становится ключевым элементом обеспечения макроэкономической устойчивости. Надёжные прогнозы инфляции необходимы не только для эффективного управления денежно-кредитной политикой, но и для повышения предсказуемости экономической динамики. Это особенно актуально для развивающихся стран, таких как Казахстан, где точные прогнозы инфляции играют важную роль в смягчении негативного влияния на покупательскую способность населения, инвестиционную активность и бюджетные решения. Таким образом, анализ инфляционных процессов имеет решающее значение для экономистов и лиц, принимающих экономические решения, стремящихся поддерживать стабильную и предсказуемую экономическую среду, способствующую устойчивому росту.

Традиционные методы прогнозирования инфляции, такие как авторегрессионные модели (Autoregressive model, AR), векторные авторегрессионные модели (Vector autoregressive model, VAR) и модели коррекции ошибок (Error correction model, ECM) (Stock and Watson, 2003), показали высокую эффективность в стабильных макроэкономических условиях. Однако, по мере усиления экономической волатильности и роста неопределённости, эти подходы сталкиваются с ограничениями. Они слабо учитывают нелинейные взаимосвязи и не адаптируются оперативно к быстрым изменениям экономической ситуации, что снижает точность прогнозов в нестабильных условиях.

В ответ на эти вызовы, современные исследования рассматривают интеграцию моделей машинного обучения (Machine learning, ML). Эти методы демонстрируют преимущества в моделировании сложных экономических процессов и улучшают точность прогнозов за счёт выявления скрытых паттернов в макроэкономических данных. Например, эмпирические исследования показывают, что алгоритмы машинного обучения могут существенно сократить ошибки прогноза по сравнению с традиционными эконометрическими моделями, особенно при высокой волатильности (Medeiros et al., 2021).

Методы машинного обучения стали важным дополнением к традиционным эконометрическим моделям, расширяя возможности анализа скрытых взаимосвязей в данных. Как отмечают Mullainathan и Spiess (2017), ML-алгоритмы позволяют гибко учитывать нелинейные зависимости и изменчивость данных, повышая адаптивность прогнозов к изменяющимся экономическим условиям. Исследования Kohlscheen (2021) и Medeiros et al. (2021) подтверждают, что ML способен эффективно обрабатывать большое количество переменных и выявлять взаимосвязи, недоступные для традиционных подходов, что способствует более полному анализу факторов, влияющих на инфляцию.

Настоящее исследование направлено на повышение точности инфляционных прогнозов путем интеграции методов машинного обучения в существующую систему селективно-комбинированного инфляционного прогноза (System of Selective Combined Inflation Forecast, SSCIF), используемую НБРК для краткосрочного прогнозирования инфляции. Ранее SSCIF успешно демонстрировала точность в краткосрочных прогнозах на основе классических эконометрических моделей. Однако современные экономические условия требуют большей гибкости и адаптивности, что обосновывает использование методов машинного обучения, таких как Ridge, Lasso и Elastic Net. Эти алгоритмы эффективно выявляют сложные взаимосвязи и скрытые паттерны, которые трудно уловить с помощью традиционных подходов (Mullainathan and Spiess, 2017; Zou and Hastie, 2005).

Основная цель данного исследования – повышение точности прогнозов инфляции за счёт интеграции методов машинного обучения с традиционными эконометрическими подходами. Такое совмещение в рамках модели SSCIF позволяет улучшить адаптацию к изменяющимся экономическим условиям и повысить качество краткосрочных прогнозов.

Исследование состоит из нескольких разделов. Во втором разделе представлен обзор литературы, а в третьем – раскрываются методологические аспекты исследования. Четвёртая часть, посвящённая результатам, включает две ключевые составляющие: первая часть рассматривает обновление коэффициентов и уравнений традиционных моделей для отражения актуальной экономической ситуации, а вторая – интеграцию трех алгоритмов машинного обучения, адаптированных к условиям экономической неопределённости. Заключительный раздел содержит выводы. Результаты исследования демонстрируют высокий потенциал методов машинного обучения для макроэкономического моделирования, улучшая точность прогнозов инфляции в Казахстане.

## **2 Обзор литературы**

На фоне мировой экономической неопределённости и ускоряющихся структурных изменений прогнозирование становится не просто аналитической задачей, но и важнейшим инструментом экономической политики. Высокая точность прогнозов играет критическую роль в поддержании макроэкономической стабильности и эффективности управленческих решений.

Прогнозирование инфляции важно для всех участников экономики, так как позволяет планировать инвестиции, расходы и сбережения в условиях неопределённости. Это особенно актуально для домохозяйств и предприятий, которые корректируют свои решения на основе ожиданий будущих цен. Для центральных банков, проводящих политику инфляционного таргетирования, точные прогнозы инфляции являются основой для принятия решений о

процентных ставках и других мерах денежно-кредитной политики. Они помогают стабилизировать инфляционные ожидания и повысить доверие к экономической политике, что критически важно для обеспечения макроэкономической устойчивости и стабильного роста.

Классические эконометрические модели занимают значимое место в прогнозировании инфляции (Stock and Watson, 2003). Эти методы показали высокую эффективность в относительно стабильной экономической среде, благодаря своей способности количественно оценивать долгосрочные и краткосрочные тенденции. Однако современные экономические условия – рост объёмов данных, высокая волатильность и глобальная неопределённость – предъявляют новые требования к гибкости и адаптивности моделей. Традиционные подходы всё чаще сталкиваются с ограничениями в способности своевременно реагировать на внешние шоки и изменения инфляционных ожиданий, что требует поиска новых решений для повышения точности прогнозов.

В условиях этих вызовов методы машинного обучения приобретают всё большее значение в экономических исследованиях. Благодаря своей гибкости, методы машинного обучения открыли новые возможности для макроэкономического прогнозирования. Mullainathan и Spiess (2017) отмечают, что такие подходы способны выявлять сложные нелинейные взаимосвязи, значительно расширяя возможности традиционных эконометрических методов. В отличие от классических моделей, алгоритмы ML адаптируются к сложным и быстро меняющимся данным, улучшая точность прогнозов. Kohlscheen (2021) и Medeiros et al. (2021) подчёркивают, что современные алгоритмы могут работать с большим количеством переменных и выявлять зависимости, которые остаются скрытыми для традиционных подходов, что позволяет глубже понять детерминанты инфляции.

Сравнительный анализ методов показывает, что методы машинного обучения обладают несколькими ключевыми преимуществами перед классическими эконометрическими подходами. Во-первых, модели ML способны обрабатывать большие массивы данных с многочисленными переменными и эффективно справляться с проблемой многомерности, что особенно важно при исследовании инфляции, зависящей от широкого спектра факторов (Kohlscheen, 2021). Во-вторых, они ориентированы на оптимизацию точности прогнозов, а не только на оценку параметров, что делает их более эффективными для краткосрочных прогнозов, необходимых для своевременного реагирования на макроэкономические изменения (Mullainathan and Spiess, 2017). В-третьих, способность моделей ML автоматически обновляться при поступлении новых данных позволяет им оперативно реагировать на изменения экономической среды и обеспечивать актуальные прогнозы в реальном времени.

В последние годы интеграция методов машинного обучения с традиционными эконометрическими моделями стала широко применяться в

прогнозировании инфляции. Такие гибридные подходы позволяют объединить интерпретируемость и учет исторических экономических тенденций, присущие эконометрическим моделям, со способностью алгоритмов машинного обучения выявлять нелинейные взаимосвязи и адаптироваться к быстро меняющимся условиям. Например, Medeiros et al. (2021) показали, что сочетание байесовских векторных авторегрессионных моделей (BVAR) с методами машинного обучения, такими как Ridge и Elastic Net регрессии, существенно снижает ошибки прогноза, особенно в периоды высокой волатильности.

Преимущество гибридных моделей заключается в их способности обеспечивать более точное прогнозирование за счет учета скрытых зависимостей, которые трудно выявить классическими методами. Kohlscheen (2021) отмечает, что включение алгоритмов, таких как адаптивный LASSO, в эконометрические модели повышает их адаптивность к изменяющимся макроэкономическим условиям и улучшает точность прогнозов.

Дальнейшие исследования подтверждают, что применение методов машинного обучения и факторных моделей существенно повышает точность прогнозов инфляции, особенно при работе с большими данными. Medeiros and Vasconcelos (2016) демонстрируют, что линейные модели с использованием метода adaLASSO повышают точность прогнозов благодаря адаптивному выбору переменных. Garcia et al. (2017) показывают, что комбинирование различных методов, таких как shrinkage-техники и регрессия на полных подмножествах, может быть особенно эффективным в экономиках с обилием данных, как это видно на примере Бразилии. Maehashi and Shintani (2020) также отмечают, что методы машинного обучения и факторные модели демонстрируют лучшую прогностическую способность при среднесрочных и долгосрочных прогнозах, что подчеркивает их универсальность.

В то же время несмотря на свои преимущества, применение методов машинного обучения сопряжено с определенными методологическими вызовами. Одним из ключевых является проблема переобучения, особенно при работе с высокоразмерными данными, характерными для макроэкономического прогнозирования. Для минимизации этого риска используются пенализованные регрессии, такие как Ridge, Lasso и Elastic Net, которые играют важную роль в эффективной структуризации и интерпретации данных (Bai and Ng, 2008; Kim and Swanson, 2018; Smeekes and Wijler, 2018). Эти методы облегчают отбор ключевых переменных и повышают устойчивость моделей, что особенно важно в условиях высокой неопределенности и волатильности.

Таким образом, гибридные модели, объединяющие методы машинного обучения и традиционные эконометрические подходы, уже занимают важное место в арсенале инструментов макроэкономического прогнозирования. Они способствуют созданию более надежных прогнозов и предоставляют центральным банкам эффективную поддержку в принятии решений.

Современные исследования подчеркивают, что сочетание машинного обучения и эконометрических методов может значительно повысить точность прогнозов инфляции, предоставляя гибкие и адаптивные модели. Однако остаются нерешенные вопросы, такие как интерпретируемость сложных моделей ML и их применимость в условиях структурных изменений в экономике. Дальнейшие исследования должны сосредоточиться на оптимизации методов машинного обучения для работы с нестационарными данными и повышении устойчивости прогнозов к внезапным экономическим шокам.

### **3 Описание данных и методология**

#### **3.1 Описание системы селективного прогнозирования инфляции и данных**

Система селективно-комбинированного прогноза инфляции представляет собой практическую реализацию подхода к селективному комбинированию прогнозов, разработанную и применяемую НБРК для краткосрочного прогнозирования инфляции в рамках системы Forecasting and Policy Analysis System (FPAS). SSCIF позволяет моделировать динамику индекса потребительских цен (ИПЦ) с использованием дезагрегированных данных, что способствует повышению точности прогнозов. В системе используются 66 субиндексов, охватывающих основные группы товаров и услуг, входящих в состав ИПЦ (Таблица 1 Приложения).

Экзогенные факторы, частично объясняющие динамику субиндексов, включают обменные курсы тенге к иностранным валютам, инфляцию в России, промышленные цены в Казахстане, объём денежной массы, реальные доходы населения, а также мировые цены на нефть и пшеницу. Этот детализированный подход позволяет более точно оценивать чувствительность инфляции к изменениям макроэкономических условий.

Все переменные в SSCIF представлены в виде сезонно-сглаженных месячных показателей, и для обеспечения стационарности временных рядов они преобразуются в разности логарифмов. Для прогнозирования используются данные, начиная с января 2011 года. SSCIF сочетает четыре типа моделей, которые подразделяются на безусловные и условные: к безусловным моделям относятся авторегрессия с линейным трендом (LTAR) и случайное блуждание (RW), а к условным – многофакторная регрессия (OLS) и байесовская векторная авторегрессия (BVAR). Точность прогнозов оценивается с помощью показателя среднеквадратической ошибки (RMSE), что позволяет взвешивать модели в зависимости от их прогностической точности.

Модель SSCIF применяет селективный подход, позволяющий оценивать точность каждой модели на различных временных горизонтах и рассчитывать комбинационные веса для отдельных временных периодов. Эти веса, которые Тулеуов (2017) называет «рекурсивными комбинационными весами», повышают точность прогнозов на ближних горизонтах и обеспечивают более гибкую адаптацию к изменяющимся условиям. Затем, на основе полученного прогноза, каждый субиндекс взвешивается согласно доле его вклада в потребительскую корзину Казахстана. Итоговый взвешенный прогноз представляет собой общий прогноз ИПЦ, который отражает прогнозируемую динамику инфляции на заданный временной период.

Таким образом, SSCIF представляет собой гибкий инструмент для прогнозирования инфляции, основанный на комбинировании моделей и учёте различных временных интервалов, что позволяет формировать более надёжные прогнозы индекса потребительских цен.

В обновленной версии модели, как и в предыдущей, основным показателем инфляции остается индекс потребительских цен. В рамках использования методов машинного обучения список используемых переменных расширился. В отличие от традиционных моделей, модели машинного обучения используют не агрегированный показатель промышленных цен, а его дезагрегированные составляющие по отдельным отраслям и товарам, что расширяет количество объясняющих переменных и позволяет более точно учитывать влияние каждого сектора экономики на инфляцию. Также в модели будут использоваться мировые цены не только на пшеницу, но и на другие важные товары, входящие в индекс продовольственных цен FAO, что позволит повысить точность прогнозов.

Добавленные переменные, как и существующие, представляют собой месячные, сезонно-сглаженные данные, трансформированные в разности логарифмов для обеспечения стационарности временных рядов на уровне значимости 5%.

### **3.2 Обновление существующих эконометрических моделей**

В рамках данного исследования в систему SSCIF были внесены значительные улучшения, направленные на повышение точности прогнозирования с использованием эконометрических моделей. Подробная информация об эконометрических моделях, используемых в SSCIF, представлена в исследовании Тулеуова (2017). Внесённые изменения в модели включают:

- 1) Улучшение уравнений и прогнозов экзогенных переменных.

Были пересмотрены уравнения для экзогенных переменных, что позволило повысить точность их прогнозов. Это улучшение обеспечивает более адекватное моделирование внешних факторов, влияющих на инфляцию в Казахстане.

## 2) Повышение точности прогнозов сезонных факторов

Модель теперь включает более точные уравнения для прогнозирования сезонных факторов, связанных с экзогенными переменными. Эти изменения позволили улучшить прогнозирование сезонных колебаний, которые играют ключевую роль в формировании инфляции, особенно в тех секторах, которые подвержены сильным сезонным изменениям.

## 3) Экзогенные прогнозы для инфляции в России и мировых цен на нефть

Прогнозирование инфляции в России и мировых цен на нефть теперь задается экзогенно в модели, основываясь на прогнозах сателлитных моделей. Этот подход позволяет учитывать внешние экономические условия с использованием более надежных и актуальных данных.

## 4) Расширение и оптимизация уравнений BVAR:

В байесовской векторной авторегрессионной модели (BVAR) количество уравнений было увеличено с 15 до 18 на основе экспертной оценки и выявленных корреляционных связей между группами товаров и услуг, входящих в состав ИПЦ. Дополнительно число лагов в модели было обновлено с использованием эконометрических тестов, таких как критерий информационного выбора (AIC) и тесты на автокорреляцию остатков, что позволило более полно учесть актуальные входные данные.

### 3.3 Использование моделей машинного обучения

В данной методологической части представлены модели машинного обучения, используемые для краткосрочного прогнозирования инфляции в Казахстане. Применение методов Ridge, Lasso и Elastic Net основано на их способности улучшать прогнозы, минимизируя такие проблемы, как мультиколлинеарность и переобучение, характерные для макроэкономических данных. Эти подходы позволяют одновременно учитывать множество экономических переменных и обеспечивать стабильность прогнозов, адаптируя модель к изменчивым экономическим условиям.

#### 3.3.1 Ridge Regression

Ridge Regression (гребневая регрессия) представляет собой метод регуляризации, широко используемый для решения проблемы мультиколлинеарности в задачах множественной линейной регрессии. Традиционные методы регрессии предполагают независимость объясняющих переменных, однако в реальности эти переменные часто оказываются взаимосвязанными. Это приводит к снижению точности оценок и делает модель нестабильной. Ridge Regression позволяет минимизировать влияние мультиколлинеарности с помощью регуляризации, добавляя штраф за величину коэффициентов регрессии, и тем самым стабилизируя модель.

Ridge Regression основана на стандартной модели множественной линейной регрессии:

$$Y = X\beta + \epsilon$$

где  $Y$  – зависимая переменная (в данном случае, инфляция),  $X$  – матрица объясняющих переменных (например, мировые цены на товары, курс валют, денежная масса),  $\beta$  – коэффициенты модели, которые нужно оценить,  $\epsilon$  – случайная ошибка.

В стандартной линейной регрессии цель состоит в минимизации суммы квадратов отклонений (RSS) для нахождения коэффициентов:

$$RSS = \sum_{i=1}^n (Y_i - X_i\beta)^2$$

Ridge регрессия добавляет к этой функции штраф за величину коэффициентов, чтобы контролировать их размер и избежать переобучения. Функция, которую минимизирует Ridge, выглядит так:

$$RSS + \lambda \sum_{j=1}^p \beta_j^2$$

где RSS – сумма квадратов отклонений,  $\lambda$  – коэффициент регуляризации, контролирующей степень "штрафа" за большие значения коэффициентов.

Этот дополнительный параметр, известный как ridge-параметр, позволяет смягчить влияние мультиколлинеарности, делая оценку модели более устойчивой и точной.

При малых значениях  $\lambda$  гребневая регрессия приближается к обычной линейной регрессии, где оценка коэффициентов производится методом наименьших квадратов. Однако по мере увеличения  $\lambda$  коэффициенты становятся меньше, что позволяет стабилизировать их значения в условиях мультиколлинеарности и уменьшить вариации модели. Это ключевая особенность Ridge Regression: она снижает влияние переменных, которые могут вносить слишком большую вариативность в модель, что позволяет улучшить точность прогнозов, особенно при использовании большого числа коррелированных переменных.

Гребневая регрессия была предложена Hoerl and Kennar (1970) и с тех пор стала одним из основных методов для работы с мультиколлинеарностью. Ridge-параметр  $\lambda$  определяет, насколько сильно модель будет "наказывать" большие значения коэффициентов, что позволяет сбалансировать точность и простоту модели. Для выбора оптимального значения  $\lambda$  часто используется метод кросс-валидации – модель проверяется на различных подвыборках данных для нахождения того значения, при котором точность прогноза максимальна.

В контексте прогнозирования инфляции Ridge Regression особенно полезна при анализе широкого набора макроэкономических переменных, таких как мировые цены на нефть, валютные курсы, денежная масса, доходы населения и другие показатели, которые часто оказываются

коррелированными. Проблема мультиколлинеарности, возникающая при наличии подобных взаимозависимых переменных, может существенно снижать точность прогнозов в традиционных моделях. Ridge Regression позволяет стабилизировать модель, обеспечивая точные прогнозы в краткосрочной и среднесрочной перспективе.

Таким образом, Ridge Regression является эффективным инструментом для построения моделей прогнозирования инфляции, способным уменьшить влияние мультиколлинеарности и переобучения. Этот метод особенно актуален при работе с макроэкономическими данными, которые характеризуются высокой сложностью и взаимосвязями между переменными.

### 3.3.2 Lasso Regression

Lasso Regression (Least Absolute Shrinkage and Selection Operator), предложенный Tibshirani (1996), представляет собой метод линейной регрессии, сочетающий регуляризацию и отбор переменных. В отличие от традиционных методов, Lasso Regression позволяет не только оценивать коэффициенты модели, но и автоматически исключать незначимые переменные, что делает его ценным инструментом для построения эконометрических моделей, в частности для прогнозирования инфляции.

Основной принцип Lasso заключается в использовании регуляризации с применением L1-нормы, добавляющей штраф за абсолютные значения коэффициентов. Это приводит к изменению стандартной функции минимизации следующим образом:

$$RSS + \lambda \sum_{j=1}^p |\beta_j|$$

где RSS - сумма квадратов отклонений,  $\lambda$  – коэффициент регуляризации.

В отличие от Ridge Regression, где штраф применяется к квадрату коэффициентов, Lasso накладывает штраф на их абсолютные значения, что позволяет уменьшать значения некоторых коэффициентов до нуля. Таким образом, Lasso не только уменьшает влияние отдельных переменных, но и исключает те, которые вносят минимальный вклад в прогноз, оставляя только значимые факторы.

Настройка значения  $\lambda$  является важным этапом в Lasso, поскольку именно он определяет количество переменных, остающихся в модели. Если  $\lambda=0$ , модель сводится к обычной линейной регрессии. При увеличении  $\lambda$ , всё больше переменных исключается из модели. Оптимальное значение  $\lambda$  подбирается с использованием кросс-валидации, чтобы достичь баланса между точностью и сложностью модели.

Lasso является мощным инструментом для прогнозирования инфляции, так как он помогает выявить ключевые факторы, влияющие на инфляционные процессы. Преимущество Lasso заключается в том, что он может автоматически исключить нерелевантные переменные, оставив только

наиболее значимые. Таким образом, Lasso полезен при работе с высокоразмерными данными, где многие переменные могут быть коррелированы между собой. В отличие от Ridge Regression, где все коэффициенты уменьшаются, Lasso может полностью исключить переменные, что упрощает интерпретацию модели и снижает риск переобучения.

### 3.3.3 Elastic Net

Метод Elastic Net используется в регрессионном анализе для решения задач, связанных с наличием большого количества объясняющих переменных, среди которых наблюдается сильная корреляция. Разработанный Zou and Hastie (2005) Elastic Net сочетает элементы двух методов регуляризации – Lasso и Ridge, что позволяет одновременно справляться с мультиколлинеарностью и производить отбор наиболее значимых переменных.

Математически Elastic Net представляет собой расширение линейной регрессии с добавлением двух штрафных функций. Основная цель этого метода заключается в минимизации следующей функции:

$$RSS + \lambda_1 \sum_{j=1}^p |\beta_j| + \lambda_2 \sum_{j=1}^p \beta_j^2$$

В этой формуле  $\lambda_1$  отвечает за L1-регуляризацию (как в Lasso), которая "обнуляет" нерелевантные коэффициенты, а  $\lambda_2$  – за L2-регуляризацию (как в Ridge), которая уменьшает величину всех коэффициентов и стабилизирует модель при наличии коррелированных переменных.

Эта комбинация L1 и L2-регуляризации делает Elastic Net особенно полезным в ситуациях, когда данные содержат множество взаимосвязанных переменных. L1-регуляризация способствует отбору переменных, исключая те, которые оказывают минимальное влияние на прогноз, а L2-регуляризация уменьшает дисперсию оценок коэффициентов, обеспечивая устойчивость модели к мультиколлинеарности. Таким образом, Elastic Net эффективно сочетает преимущества Ridge и Lasso, обеспечивая более надежные результаты при наличии сложных и многомерных данных. Одним из ключевых этапов в Elastic Net является выбор оптимальных значений для параметров  $\lambda_1$  и  $\lambda_2$ , что обычно делается с помощью кросс-валидации. Этот процесс помогает найти баланс между отбором переменных и стабилизацией коэффициентов, что улучшает точность прогноза.

## 4 Результаты

### 4.1 Улучшение эконометрических моделей

Внесённые изменения в эконометрическую модель SSCIF значительно повысили её прогнозную эффективность, снизив ошибки прогнозирования и улучшив точность оценки ключевых показателей. Это позволило более эффективно учитывать взаимосвязи между экономическими переменными, что повышает надёжность полученных результатов.

В Таблице 1 представлена количественная оценка точности псевдо-вневыборочных прогнозов инфляции в Казахстане для исходной и обновлённой версий модели SSCIF<sup>3</sup>. Точность прогнозов измерялась с использованием показателя среднеквадратичной ошибки (RMSE), где более низкие значения свидетельствуют о повышении качества прогнозирования модели. Согласно представленным данным, обновлённая модель демонстрирует улучшение точности прогнозов для большинства периодов. Однако в период с марта по ноябрь 2016 года предыдущая модель показала меньшую ошибку RMSE в прогнозировании непродовольственной и сервисной инфляции, что связано с тем, что она изначально была разработана с учётом особенностей этого временного интервала. Данные обновлённой модели, напротив, показывают более высокие ошибки в прогнозах для данного периода, что связано с её адаптацией к более широкой временной выборке и учётом новых данных, полученных в последующие годы.

Для периода с января по сентябрь 2018 года обновлённая модель продемонстрировала существенное снижение среднеквадратичной ошибки (RMSE) для всех категорий инфляции. Наибольшее улучшение наблюдается для общей инфляции, где ошибка сократилась на 18,6%. Эти результаты подчёркивают способность обновлённой модели учитывать новые экономические тренды, которые не были отражены в предыдущей версии.

В период с июля 2020 года по март 2021 года обновлённая модель также показала значительные улучшения, особенно для общей инфляции, где ошибка прогноза снизилась на 15,8%. Это говорит о том, что обновления позволили учитывать более долгосрочные взаимосвязи между экономическими переменными. Для последнего периода с мая 2023 года по январь 2024 года обновлённая модель снизила RMSE для общей инфляции на 15,3%. Это свидетельствует о том, что модель эффективно адаптировалась к новым макроэкономическим условиям, включая изменения в динамике цен и курсовые факторы.

Таким образом, предыдущая модель демонстрирует лучшую точность для ранних периодов, таких как март – ноябрь 2016 года, благодаря своему

---

<sup>3</sup> Прогнозы, представленные в данном исследовании, являются исключительно модельными оценками, основанными на внутренних алгоритмах эконометрических моделей. Они не включают дополнительные факторы, такие как внешние предпосылки, экспертные суждения, геополитическая ситуация, изменения в регулируемых ценах и другие административные меры. Поэтому эти результаты могут отличаться от реальных прогнозов, которые учитывают широкий спектр внешних факторов и экспертных оценок, влияющих на экономическую ситуацию и динамику цен.

изначальному фокусу на этих временных интервалах. Однако для более поздних периодов обновлённая модель существенно превосходит предыдущую, так как она учитывает больший объём данных и адаптирована к новым экономическим условиям.

На рисунках 1–4 показано сравнение фактических данных и псевдо-вневыборочных прогнозов инфляции для обеих версий модели SSCIF. Обновлённая модель немного точнее отражает месячные и годовые изменения, подтверждая её улучшенную способность учитывать экономические изменения и прогнозировать компоненты ИПЦ.

Таким образом, обновленная версия оригинального SSCIF обеспечивает более высокий уровень точности прогнозов инфляции, что делает её более надёжным инструментом для прогнозирования инфляции.

**Таблица 1. Количественная оценка точности псевдо-вневыборочных прогнозов инфляции (ИПЦ в % м/м) в Казахстане на различных тестовых участках фактической выборки для предыдущей и обновленной модели**

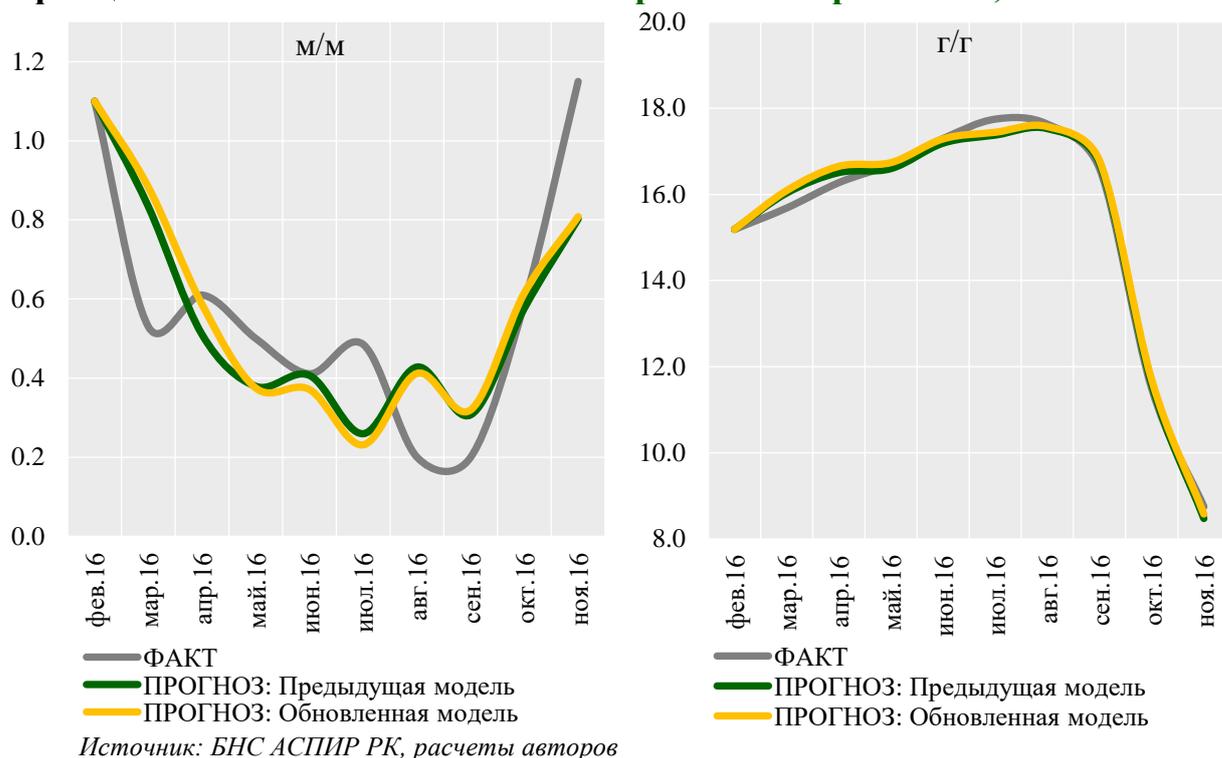
Период		Продовольственная инфляция	Непродовольственная инфляция	Сервисная инфляция	Общая инфляция
<b>Март 2016 г. - Ноябрь 2016 г.</b>	Предыдущая модель (RMSE)	0.519	0.210	0.281	<b>0.198</b>
	Обновленная модель (RMSE)	0.494	0.249	0.285	0.206
<b>Январь 2018 г. - Сентябрь 2018 г.</b>	Предыдущая модель (RMSE)	0.331	0.627	0.277	0.242
	Обновленная модель (RMSE)	0.312	0.543	0.247	<b>0.197</b>
<b>Июль 2020 г. - Март 2021 г.</b>	Предыдущая модель (RMSE)	0.124	0.323	0.293	0.076
	Обновленная модель (RMSE)	0.134	0.263	0.290	<b>0.064</b>
<b>Май 2023 г. - Январь 2024 г.</b>	Предыдущая модель (RMSE)	0.484	0.347	0.184	0.203
	Обновленная модель (RMSE)	0.464	0.293	0.184	<b>0.172</b>

Источник: БНС АСПИР РК, расчеты авторов

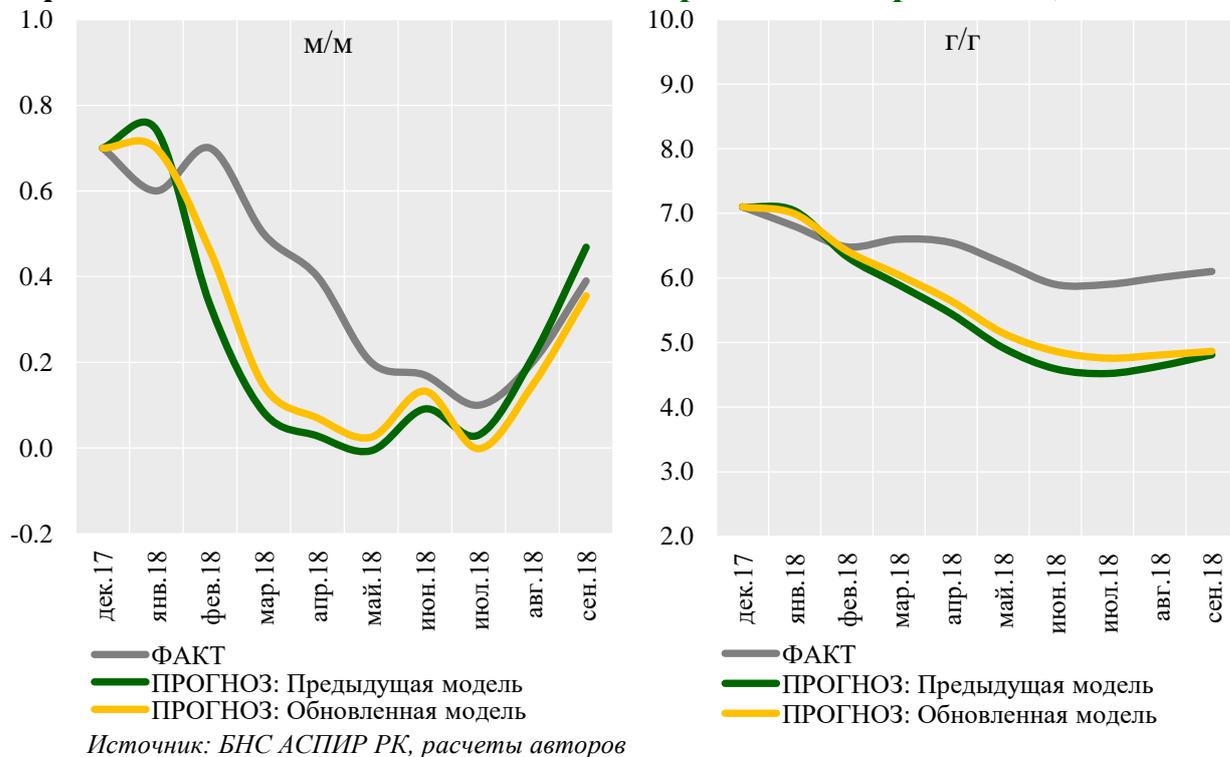
Цветовая шкала точности



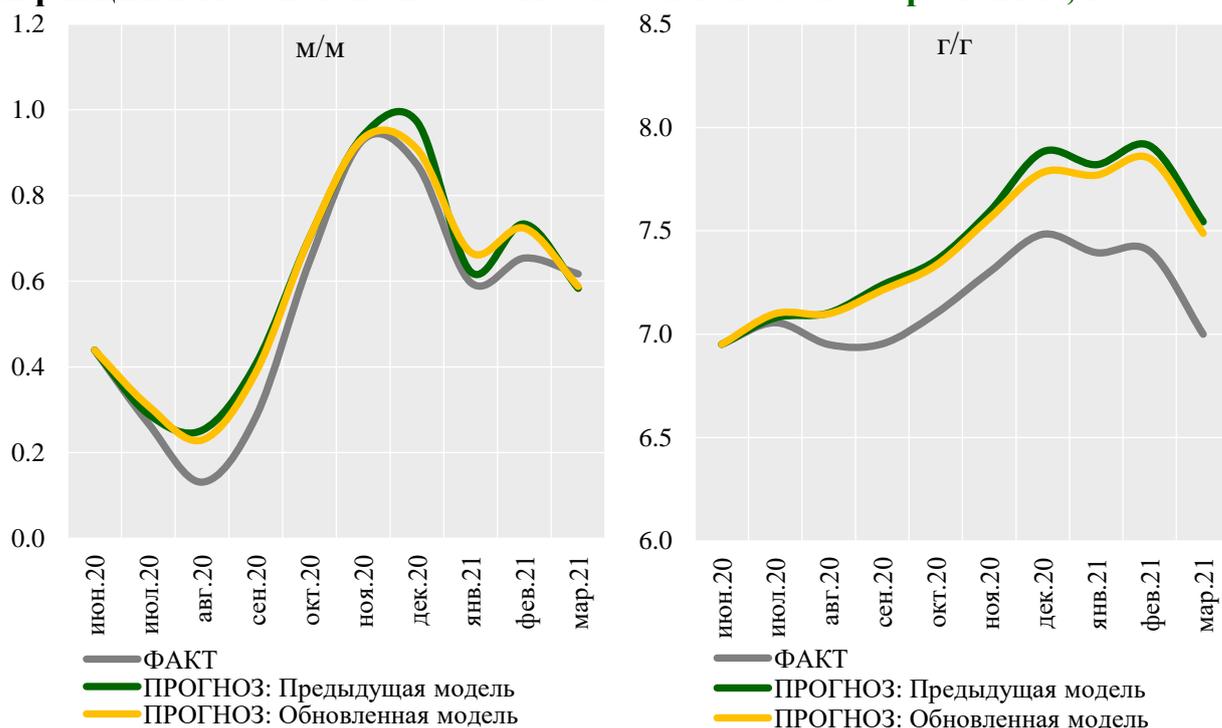
**Рисунок 1. Сравнение факта и реального вне-выборочного прогноза инфляции в Казахстане по SSCIF с марта по ноябрь 2016 г., в %**



**Рисунок 2. Сравнение факта и реального вне-выборочного прогноза инфляции в Казахстане по SSCIF с января по сентябрь 2018 г., в %**

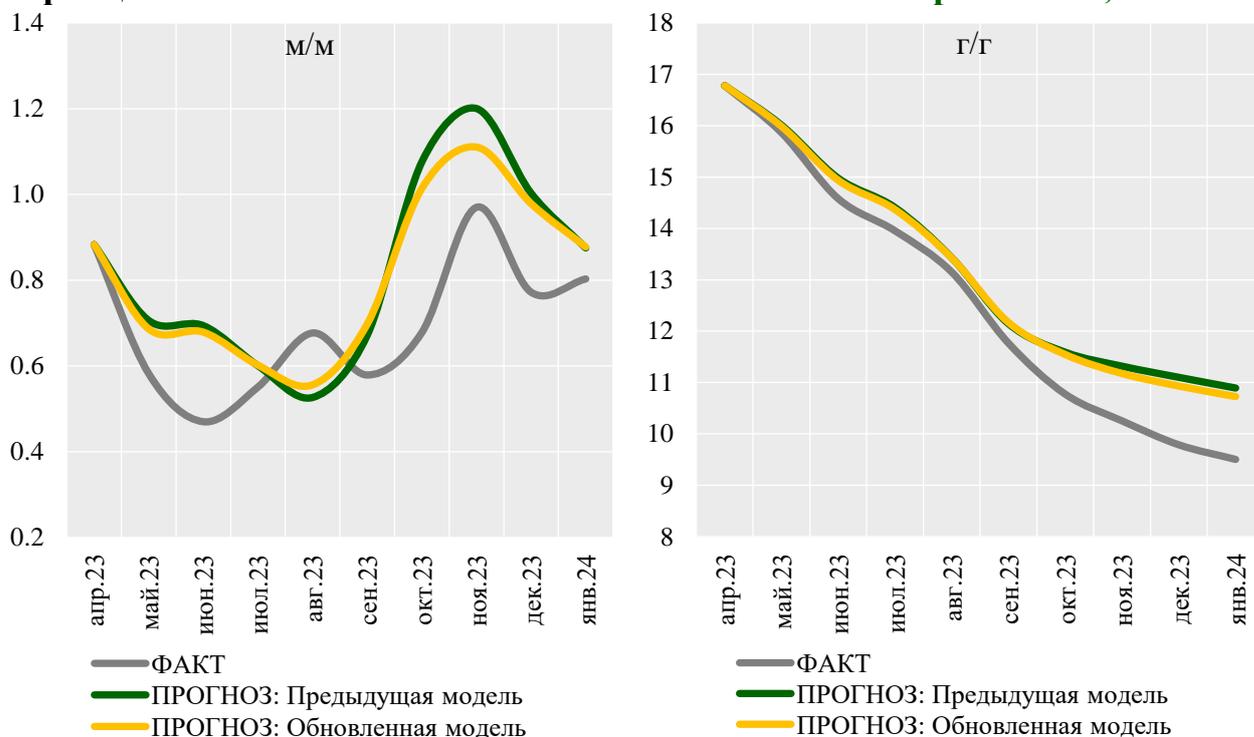


**Рисунок 3. Сравнение факта и реального вне-выборочного прогноза инфляции в Казахстане по SSCIF с июля 2020 г. по март 2021 г., в %**



Источник: БНС АСПИР РК, расчеты авторов

**Рисунок 4. Сравнение факта и реального вне-выборочного прогноза инфляции в Казахстане по SSCIF с мая 2023 г. по январь 2024 г., в %**



Источник: БНС АСПИР РК, расчеты авторов

## 4.2 Использование моделей машинного обучения

Результаты количественной оценки точности псевдо-вне-выборочных прогнозов инфляции, представленные в Таблице 2, демонстрируют различия в эффективности традиционных моделей и моделей машинного обучения (ML) на различных временных интервалах.

В период с марта по ноябрь 2016 года ML-модель оказалась менее эффективной в данном интервале, особенно в прогнозировании непродовольственной и общей инфляции. Это объясняется тем, что она оптимизирована для более поздних временных периодов и использует более широкий набор данных.

В период с января по сентябрь 2018 года обновлённая модель показала более высокую точность в прогнозировании продовольственной и сервисной инфляции, что подтверждается меньшей среднеквадратичной ошибкой (RMSE). Однако расширенная модель с использованием методов ML превзошла обновлённую модель в прогнозировании непродовольственной и общей инфляции, демонстрируя более высокую адаптивность к изменениям в экономической среде. Снижение RMSE для общей инфляции в этом периоде составило 12,2%.

**Таблица 2. Количественная оценка точности псевдо-вне-выборочных прогнозов инфляции (ИПЦ в % м/м) в Казахстане на различных тестовых участках фактической выборки для обновленной и расширенной модели**

Период		Продовольственн ая инфляция	Непродовольств енная инфляция	Сервисная инфляция	Общая инфляция
<b>Март 2016 г. - Ноябрь 2016 г.</b>	Обновленная модель (RMSE)	0.494	0.249	0.285	<b>0.206</b>
	Расширенная модель с ML (RMSE)	0.771	0.662	0.336	0.473
<b>Январь 2018 г. - Сентябрь 2018 г.</b>	Обновленная модель (RMSE)	0.312	0.543	0.247	0.197
	Расширенная модель с ML (RMSE)	0.340	0.351	0.265	<b>0.173</b>
<b>Июль 2020 г. - Март 2021 г.</b>	Обновленная модель(RMSE)	0.134	0.263	0.290	0.064
	Расширенная модель с ML (RMSE)	0.221	0.178	0.258	<b>0.060</b>
<b>Май 2023 г. - Январь 2024 г.</b>	Обновленная модель(RMSE)	0.464	0.293	0.184	0.172
	Расширенная модель с ML(RMSE)	0.276	0.167	0.234	<b>0.131</b>

Источник: БНС АСПИР РК, расчеты авторов

Цветовая шкала точности

менее точно

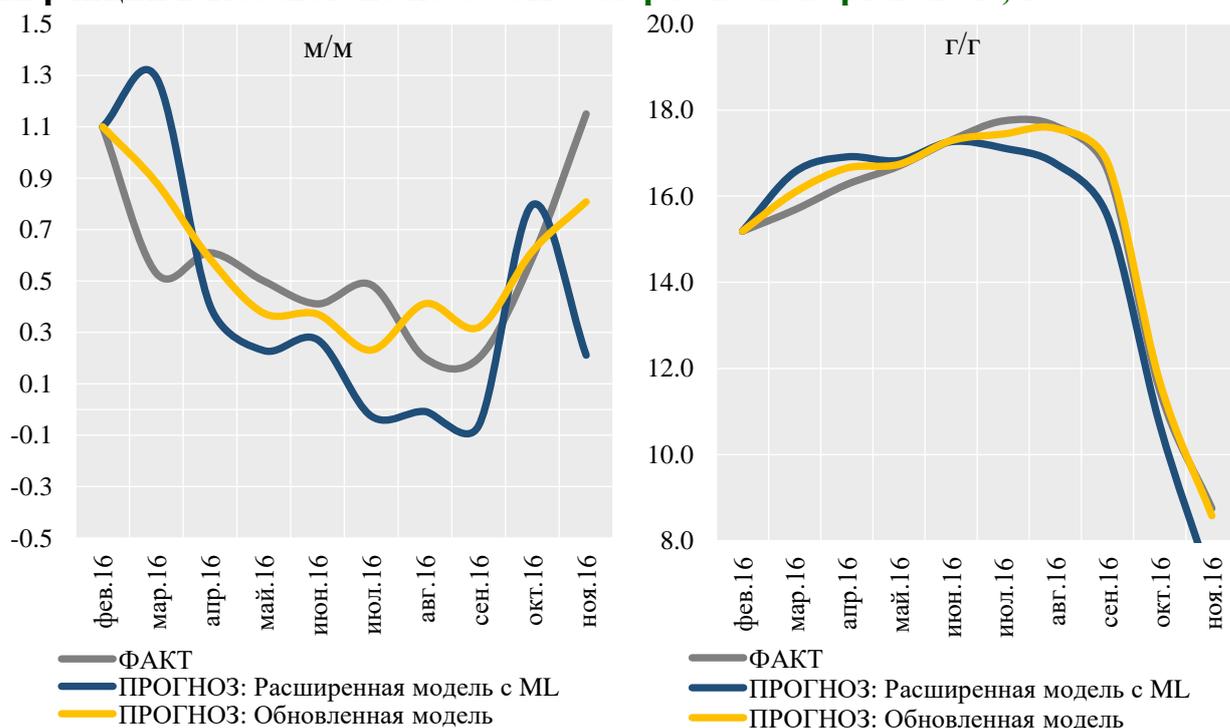
более точно

В период с июля 2020 года по март 2021 года ML-модель продемонстрировала значительное улучшение в прогнозировании непродовольственной и общей инфляции, снизив RMSE на 32,3% и 6,3% соответственно, по сравнению с обновлённой моделью. Тем не менее, для продовольственной и сервисной инфляции обновлённая модель сохранила лучшие результаты, показывая стабильность в прогнозировании этих компонентов инфляции.

На более позднем временном интервале, охватывающем период с мая 2023 по январь 2024 года, ML-модель продемонстрировала явное преимущество в прогнозировании практически всех категорий инфляции. Особенно заметно снижение RMSE для непродовольственной инфляции, что указывает на высокую точность и гибкость ML-подхода в условиях современных экономических вызовов и нестабильности. Тем не менее, в сегменте сервисной инфляции обновлённая модель сохранила лидерство, обеспечивая более низкую ошибку прогноза. Это подтверждает, что традиционные подходы остаются актуальными для прогнозирования более стабильных и медленно изменяющихся компонентов инфляции. Для общей инфляции ML-модель обеспечила снижение RMSE на 23,8%, что демонстрирует её адаптивность в условиях новых экономических вызовов.

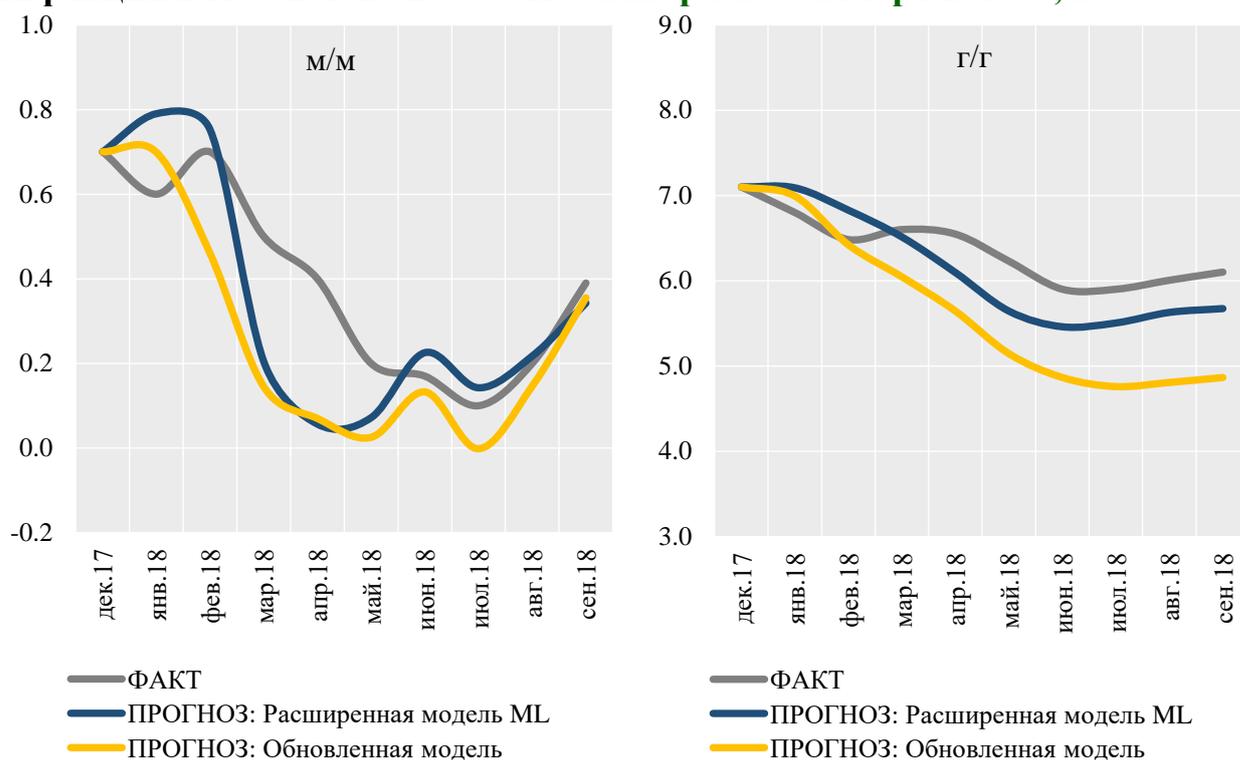
Использование методов машинного обучения позволяет существенно повысить точность прогнозирования инфляции (Рисунки 5-8). Однако традиционные подходы, заложенные в обновлённую модель, остаются эффективными для некоторых сегментов, таких как сервисная инфляция, и могут дополнять ML-модели в рамках комплексного подхода к прогнозированию.

**Рисунок 5. Сравнение факта и реального вне-выборочного прогноза инфляции в Казахстане по SSCIF с марта по ноябрь 2016 г., в %**



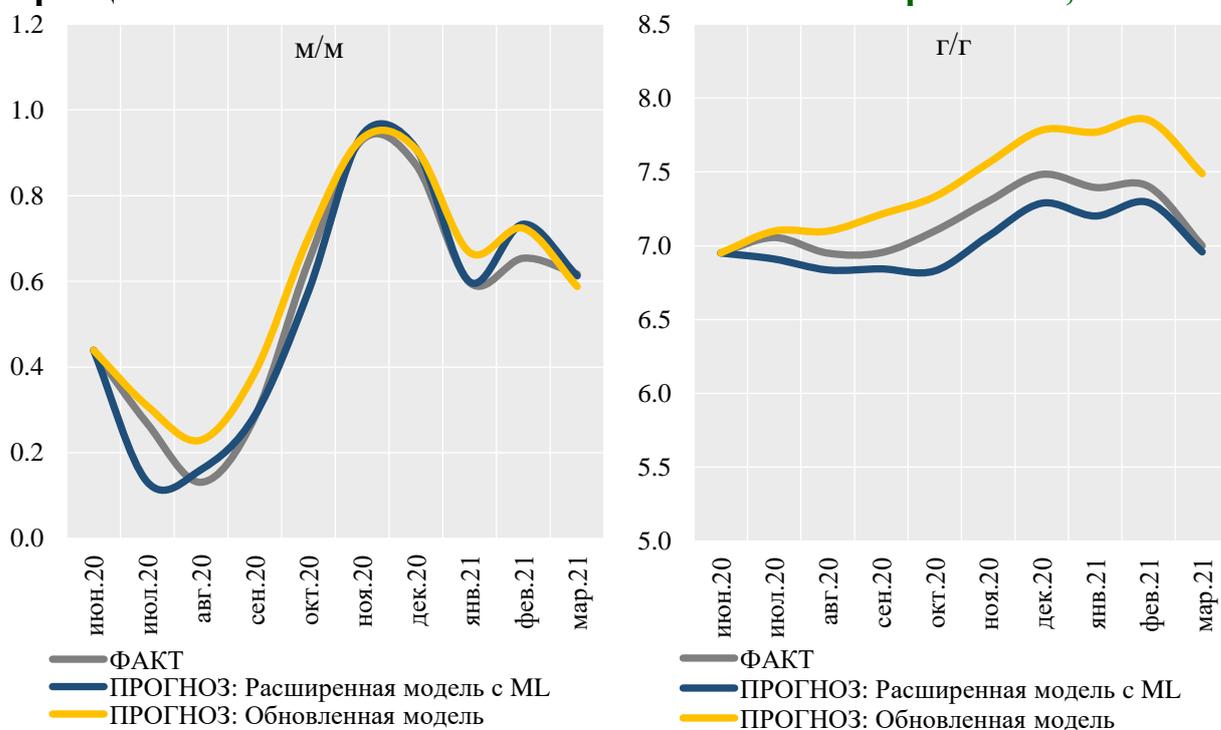
Источник: БНС АСПИР РК, расчеты авторов

**Рисунок 6. Сравнение факта и реального вне-выборочного прогноза инфляции в Казахстане по SSCIF с января по сентябрь 2018 г., в %**



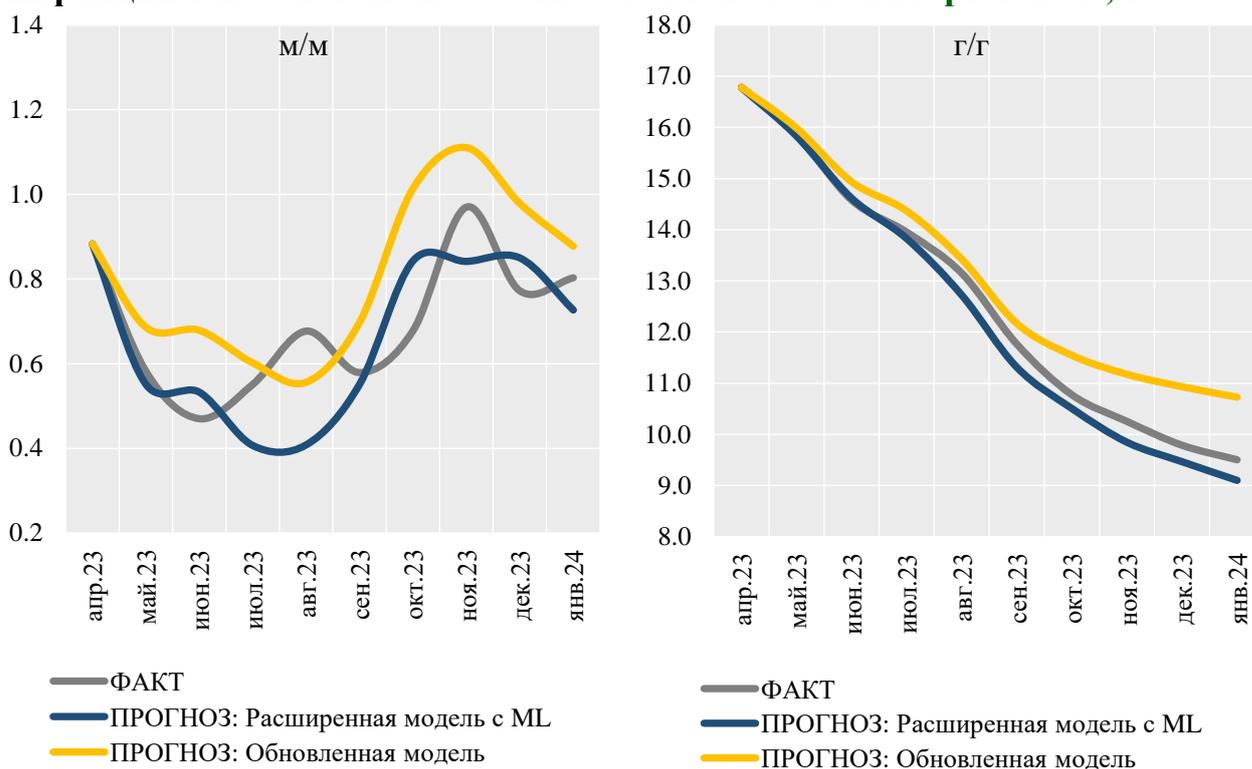
Источник: БНС АСПИР РК, расчеты авторов

**Рисунок 7. Сравнение факта и реального вне-выборочного прогноза инфляции в Казахстане по SSCIF с июля 2020 г. по март 2021 г., в %**



Источник: БНС АСПИР РК, расчеты авторов

**Рисунок 8. Сравнение факта и реального вне-выборочного прогноза инфляции в Казахстане по SSCIF с мая 2023 г. по январь 2024 г., в %**



Источник: БНС АСПИР РК, расчеты авторов

Таким образом, традиционные модели сохраняют свои преимущества в прогнозировании стабильных компонентов инфляции, особенно на ранних временных интервалах. В то же время ML-модели показывают значительное улучшение прогнозной точности в поздних периодах, что связано с их способностью учитывать широкий спектр экономических факторов, адаптироваться к изменяющимся условиям и использовать более обширную выборку данных для обучения.

## 5 Выводы

Настоящее исследование продемонстрировало высокий потенциал интеграции методов машинного обучения (ML) в систему селективно-комбинированного прогнозирования инфляции (SSCIF) для повышения точности краткосрочных прогнозов инфляции в Казахстане. Внесённые изменения в модель SSCIF и внедрение ML-алгоритмов существенно улучшили её прогнозную эффективность, снизив среднеквадратичную ошибку (RMSE) и повысив надёжность оценки ключевых показателей за счёт более эффективного учёта взаимосвязей между экономическими переменными.

Анализ результатов показал, что гибридные модели, объединяющие традиционные эконометрические подходы и современные ML-алгоритмы, способны значительно снижать ошибки прогноза (RMSE), особенно в условиях макроэкономической нестабильности и высокой волатильности. В

частности, добавление методов ML, таких как Ridge Regression, Lasso Regression и Elastic Net, в SSCIF продемонстрировало превосходство над обновлённой эконометрической моделью в прогнозировании продовольственной, непродовольственной и общей инфляции на более поздних тестовых участках. Эти алгоритмы обладают высокой адаптивностью и способностью выявлять скрытые нелинейные взаимосвязи в макроэкономических данных.

В то же время традиционные эконометрические подходы остаются эффективными для прогнозирования сервисной инфляции, что подчёркивает их ценность как инструмента анализа стабильных и медленно изменяющихся компонентов инфляции.

Интеграция ML-алгоритмов позволила учитывать более широкий спектр макроэкономических факторов, включая дезагрегированные промышленные цены и глобальные товарные индексы, что существенно расширило аналитические возможности модели. При этом сохранение элементов традиционных методов, таких как байесовская векторная авторегрессия (BVAR), остаётся важным для обеспечения интерпретируемости и устойчивости моделей в условиях ограниченных данных.

Полученные результаты имеют важное практическое значение для разработки инструментов макроэкономического прогнозирования и повышения эффективности денежно-кредитной политики. Гибридные подходы могут способствовать улучшению качества принимаемых решений, что особенно актуально для экономик с высокой степенью неопределённости, таких как Казахстан. Комплексное использование ML-подходов и традиционных эконометрических моделей, таких как обновлённая SSCIF, обеспечивает наиболее сбалансированный и эффективный инструмент прогнозирования, позволяя учитывать преимущества каждого подхода: адаптивность и гибкость ML-моделей и устойчивость традиционных методов.

Результаты исследования подчёркивают необходимость дальнейшего развития и интеграции ML-методов в экономический анализ и прогнозирование. Будущие исследования должны быть направлены на оптимизацию интеграции ML-методов в макроэкономическое прогнозирование, с акцентом на повышение интерпретируемости моделей, адаптацию к нестационарным данным и минимизацию риска переобучения. Это позволит повысить качество экономических прогнозов и эффективность принятия решений, что особенно важно в условиях высокой неопределённости и нестабильности мировой экономики.

## 6 Список использованной литературы

1. Bai, J., & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146, 304–317.
2. Garcia, M.G.P., Medeiros, M.C., & Vasconcelos, G.F.R. (2017). Real-time inflation forecasting with high-dimensional models: The case of Brazil. *International Journal of Forecasting*, 33(3), 679-693.
3. Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67.
4. Kim, H. H., & Swanson, N. R. (2018). Mining big data using parsimonious factor, machine learning, variable selection and shrinkage methods. *International Journal of Forecasting*, 34(2), 339-354.
5. Kohlscheen, E. (2021). What does machine learning say about the drivers of inflation? BIS Working Papers 980.
6. Maehashi, K., & Shintani, M. (2020). Macroeconomic forecasting using factor models and machine learning: an application to Japan. *Journal of the Japanese and International Economies*, 58, 101104.
7. Medeiros, M. C., Vasconcelos, G. F., Veiga, A., & Zilberman, E. (2021). Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *Journal of Business & Economic Statistics*, 39(1), 98-119.
8. Medeiros, M. C., & Vasconcelos, G. F. R. (2016). Forecasting macroeconomic variables in data-rich environments. *Economics Letters*, 138, 50-52.
9. Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87-106.
10. Smeekes, S., & Wijler, E. (2018). Macroeconomic forecasting using penalized regression methods. *International Journal of Forecasting*, 34(3), 408-430.
11. Stock, J. H., & Watson, M. W. (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, 41(3), 788-829.
12. Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.
13. Tuleuov, O. (2017). System of selective-combined inflation forecast (SSCIF): Choosing the optimal technique for forecasting the dynamics of consumer prices in the context of a structural shock (case of Kazakhstan) // Тулеуов, О. (2017). Система селективно-комбинированного прогноза инфляции (SSCIF): выбор оптимальной техники прогнозирования динамики потребительских цен в условиях структурного шока (на примере Казахстана) (No. 2017-9).
14. Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2), 301-320.

## 7 Приложение

Таблица 1. Исходные данные SSCIF

Субиндексы цен на продовольственные товары (13 субиндексов)	Субиндексы цен на непродовольственные товары (23 субиндекса)	Субиндексы цен на услуги (30 субиндексов)	Экзогенные переменные
Хлебобулочные изделия и крупы	Материалы для изготовления одежды	Чистка, ремонт и прокат одежды	Цена на нефть марки Brent
Мясо	Верхняя одежда	Ремонт и прокат обуви	Номинальный обменный курс USD/KZT
Рыба и морепродукты	Другие предметы одежды и аксессуары одежды	Фактическая арендная плата за жилье	Номинальный обменный курс RUB/KZT
Молочные изделия, сыр и яйца	Ботинки, туфли и прочая обувь	Услуги по обслуживанию и ремонту жилых помещений	Номинальный обменный курс USD/RUB
Масла и жиры	Материалы для обслуживания и ремонта жилых помещений	Водоснабжение	Продовольственная инфляция в России
Фрукты	Твердое топливо	Сбор мусора	Непродовольственная инфляция в России
Овощи	Мебель, предметы домашнего обихода, ковры и другие покрытия для пола, их ремонт	Канализация	Сервисная инфляция в России
Сахар, джем, мед, шоколад и кондитерские изделия	Текстильные изделия, используемые в домашнем хозяйстве	Прочие услуги, связанные с содержанием жилых помещений	Цены в промышленности в Казахстане
Продукты питания, не отнесенные к другим категориям	Бытовые приборы	Электроэнергия	Цены в обрабатывающей промышленности в Казахстане
Кофе, чай и какао	Стекланые изделия, столовые приборы и домашняя утварь	Газ	Тенговая денежная масса в Казахстане
Минеральная вода, прохладительные напитки, фруктовые и овощные соки	Инструменты и приспособления, используемые в быту и садоводстве	Тепловая энергия	Реальный денежный доход в Казахстане
Алкогольные напитки	Товары и услуги, используемые для ведения домашнего хозяйства	Амбулаторные услуги	Индекс FAO Cereals
Табачные изделия	Медикаменты, лечебное оборудование и аппаратура	Услуги больниц	
	Покупка автотранспортных средств	Техническое обслуживание и ремонт личных транспортных средств	
	Запасные части и принадлежности для личных транспортных средств	Прочие услуги, связанные с личными транспортными средствами	
	Горюче-смазочные материалы для личных транспортных средств	Услуги транспорта	
	Аудиовизуальное оборудование и фотоаппаратура, оборудование для обработки информации	Связь	
	Другие крупные товары длительного пользования для организации отдыха и культурных мероприятий	Услуги в области отдыха, развлечений и культуры	
	Другие товары и оборудование для отдыха, спорта, садоводства и домашние животные	Организация комплексного отдыха	
	Газеты, книги и канцелярские товары	Дошкольное и начальное образование	
	Электрические приборы личного пользования	Среднее образование	
	Прочие предметы, приборы и товары личного пользования	Продолженное среднее образование	
	Предметы личного ухода, не отнесенные к другим категориям	Высшее образование	
		Образование, не подразделенное по ступеням	
		Услуги общественного питания	
		Гостиничное обслуживание	
		Услуги парикмахерских и заведений личного обслуживания	
		Страхование	
		Финансовые услуги, не отнесенные к другим категориям	
		Прочие услуги, не отнесенные к другим категориям	

Источник: составлено Тулеуов (2017) и авторами по данным БНС АСПР РК, Национального Банка Казахстана, Федеральной службы по государственной статистике России