

Прикладная эконометрика, 2025, т. 79, с. 27–49.

Applied Econometrics, 2025, v. 79, pp. 27–49.

DOI: 10.22394/1993-7601-2025-79-27-49

Н.М. Макеева¹

Влияние версии ревизии официальной статистики на точность моделей наукастинга макроэкономических показателей России

В работе представлены результаты анализа точности моделей наукастинга для ВВП России и его компонентов по использованию за период с 1 квартала 2014 г. по 3 квартал 2023 г. Новизна исследования заключается в сопоставлении точности целого спектра моделей (MIDAS-, MFBVAR-, DFM-моделей, моделей с регуляризацией, а также классической авторегрессии первого порядка), оцененных на первой и финальной версиях ревизии официальной статистики с учетом доступности данных внутри рассматриваемого квартала. Сделаны выводы о стратегии оптимального подбора данных и моделей для наиболее точного наукастинга первой и финальной версий ревизии официальной статистики для ВВП и его компонентов по использованию. Также показано, что модель наукастинга для ВВП России, оцененная на первой ревизии, оказывается более точной при любой полноте данных внутри квартала, несмотря на многократный последующий пересмотр первоначальных данных Росстатом. Более того, для моделей ВВП точнее всего оказываются прогнозы первой, самой шумной и неточной, ревизии.

Ключевые слова: наукастинг; ВВП России; временные ряды; краткосрочное прогнозирование; пересмотр данных.

JEL classification: C53; E37.

1. Введение

Органы государственной власти и частный сектор заинтересованы в получении информации о текущей ситуации в экономике. Важность получаемых данных особенно возрастает в периоды кризисов и нестабильности, что усиливается двумя основными особенностями, связанными с публикацией официальной статистики. Первая проблема заключается в значительной задержке публикации данных за отчетный период: разрыв между окончанием отчетного периода и первой публикацией официальной статистики может достигать до трех месяцев. Второй, не менее важной проблемой является корректировка официальной статистики, вызванная несовершенством и последующими многократными уточнениями исходных данных. По данной причине корректировки, например, темпов роста экономики, могут проходить более четырех раз по мере обновления данных, используемых для расчетов.

¹ Макеева Наталья Михайловна — Национальный исследовательский университет «Высшая школа экономики», Москва; nmakeeva@hse.ru.

Проблема задержки официальной статистики успешно решается через модели наукастинга — оценку текущего макроэкономического состояния на основе уже вышедшей статистики за рассматриваемый период. Например, в работе (Evans, 2005) предлагается авторская модель на основе ежедневных временных рядов, позволяющая получать оценку основных макроэкономических показателей (ВВП, инфляция, безработица и т. д.) в режиме реального времени. Вопрос повышающегося уровня неопределенности в околокризисные периоды и одновременно возрастающей роли способа оперативной оценки макроэкономической ситуации на примере Великобритании в 2007–2008 гг. рассматривается в исследовании (Galvão, Mitchell, 2019). Важным результатом работы стала методика определения степени макроэкономической неопределенности для последующего использования в прогнозировании. В (Clarida et al, 2002) предлагается использовать RTFS-модель как способ агрегации месячных показателей для расчета показателя, аналогичного квартальным темпам роста реального ВВП США. Авторы говорят о простоте понимания и поддержания метода, а также об оперативности получения интересующих показателей за текущий период, что подтверждает перспективность предлагаемого подхода.

Исследование (Howrey, 1978) поднимает вопрос игнорирования статистической разницы между предварительными и конечными показателями официальной статистики. Автор демонстрирует, что использование предварительной, самой первой официальной статистики дает ложную волатильность в построении прогнозов, и предлагает метод коррекции данных для минимизации ошибки получаемого прогноза. В более поздней работе (Howrey, 1984) изучается степень связи величины ошибки прогноза инвестиций в товарно-материальные запасы с недостатками первых версий статистики. Автор отмечает, что достичь наиболее точного прогноза при прочих равных условиях можно за счет оптимального использования предварительных и пересмотренных данных. В работе (Clements, Galvão, 2012) рассматривается многомерная VAR-модель с использованием прошлых ревизий официальной статистики для получения оперативной оценки уровня выпуска и инфляции в реальном времени как для будущей ревизии, так и для финального показателя. Логическим продолжением развития инструментария является статья (Cartiero et al., 2015), где авторы предлагают модель на основе байесовской VAR-модели, которая обеспечивает более точные прогнозы инфляции после пересмотра статистики, чем другие модели, представленные в литературе. Схожая методология применяется также в (Kishor, Koenig, 2012). Метод Хауэри применяется для учета изменений в версиях официальной статистики на примере данных о занятости и уровне безработицы, реальном ВВП и уровне безработицы, а также реальном ВВП и соотношении ВВП/потребление. Авторы демонстрируют, что для каждого из рассмотренных кейсов предлагаемая процедура оценки макроэкономических показателей превосходит как традиционную VAR-модель, так и классические методы решения проблемы пересмотра данных, встречающиеся в существующей литературе.

Важность результатов пересмотра статистики на примере инфляции поднимается в работе (Croushore, 2008), где отмечается, что итоги пересмотра инфляции играют важную роль в анализе результирующего показателя федеральной инфляции, на основе которого принимаются решения по денежно-кредитной политике. Исследование (Galvão, 2017) рассматривает проблему пересмотра данных, используемых для расчета макропоказателей, в DSGE-моделях. В качестве решения предлагается подход, основанный на использовании всех версий пересмотренных данных до момента первой публикации целевой переменной, что повышает точность прогноза. Работа (Jacobs, van Norden, 2011) посвящена проблеме

краткосрочного прогнозирования версий ревизий официальной статистики. В статье описывается модель пространства состояний, учитывающая различные типы погрешностей в данных, включая шумовые и новостные эффекты, возникающие при пересмотре статистики. В (Rossi, Sekhposyan, 2015) исследуется вопрос распределения ошибок прогнозирования при построении наукастов и краткосрочных прогнозов в зависимости от различных условий. Авторы формируют индекс неопределенности и сравнивают его с другими аналогами для оценки эффективности предлагаемого метода.

Работа (Aastveit et al., 2014) предлагает подход, основанный на методе комбинированных прогнозов, для наукаста ВВП США в реальном времени. На данных со 2 квартала 1990 г. по 3 квартал 2010 г. авторы демонстрируют, что точность оперативного прогноза улучшается по мере использования наиболее свежей ревизии статистики с последующим включением в модель. В статье (Augoba, 2008) доказывается, что изменения в ревизиях официальной статистики не являются величиной с нулевым средним, что может свидетельствовать о предвзятости статистических агентств при первоначальных публикациях. Clements, Galvão (2017) предлагают использовать опросы для прогнозирования второй оценки роста ВВП, при этом наиболее точную (третью) оценку ВВП дают модели с использованием ежедневных данных о доходности фондового рынка.

Работа (Hayashi, Tachi, 2023) оценивает темпы роста ВВП Японии на данных в реальном времени и использует новый метод для анализа пересмотра официальной статистики, связанный как с появлением новых наблюдений, так и с пересмотром исходных данных. Chikamatsu et al. (2018) освещают проблему наукастинга ВВП Японии, применяя целый набор прогнозных моделей — различные вариации связующих уравнений, MIDAS- и FA-MIDAS-модели. Результатом работы является сравнение точности используемых моделей, а также различных комбинаций наукастов. Исследование (Lehmann, 2023) посвящено созданию в реальном времени базы данных для региональных экономических счетов Германии, содержащей актуальную информацию о девяти макроэкономических показателях в 16 федеральных землях Германии. Проведена оценка точности краткосрочных прогнозов и сделан вывод о влиянии корректировок первых оценок официальной статистики на результаты наукастов. В работе (Dahlhaus et al., 2017) DFM-модели применяются для точного и своевременного определения текущего роста ВВП в Бразилии, России, Индии, Китае и Мексике, с учетом особенностей в публикации данных. Авторы показывают, что драйверами изменения прогнозов являются страновые показатели, в то время как внешние переменные играют второстепенную роль. В исследовании (Kočenda, Poghosyan, 2020) рассматривается эффективность различных моделей прогнозирования для стран Европейского Союза с использованием жестких и мягких показателей. Результаты показали, что ретроспективные и краткосрочные прогнозы хорошо работают для новых стран ЕС, несмотря на большую вариативность их данных по сравнению со старыми странами ЕС, т.к. экономическая значимость прогнозирования в среднем выше.

Статья (Serpi et al., 2020) оценивает прогностическую способность скрытых факторов экономической политики и факторов неожиданности для прогнозирования ВВП с использованием эконометрических факторных моделей. Авторы показывают, что применение LASSO регуляризации позволяет получить лучший по точности прогноз темпов роста ВВП в развивающихся рыночных экономиках. Метод, предлагаемый в (Giovannelli et al., 2020), позволяет с помощью месячных индикаторов построить краткосрочный прогноз для 16 ключевых компонентов ВВП, а достаточно высокая точность наукаста достигается за счет комбинации индивидуальных прогнозов на основе экономических индикаторов. Статья (Hayashi,

Tachi, 2021) посвящена исследованию изменения наукаста в зависимости от пересмотра исходных данных и выхода новых. Результатом работы было выявление вклада отдельных объясняющих переменных в изменения полученного наукаста.

Исследование (Strohsal, Wolf, 2020) рассматривает вопрос пересмотра данных в национальных счетах Германии. Авторы отмечают, что пересмотр данных национальных счетов является необъективным, масштабным и предсказуемым, и предлагают использовать метод фильтрации исходных данных, позволяющий улучшить точность наукаста на 23%. Основная идея статьи (Anesti et al., 2022) заключается в предложении динамической факторной модели, дополненной статистическими данными (RA-DFM-модель), которая позволяет количественно оценить роль страновых данных в прогнозировании как первых оценок ВВП, так и последующих пересмотров официальных оценок. В работе показано, что RA-DFM-модель с использованием экономических данных дает более точный наукаст по сравнению со стандартной DFM-моделью при прогнозировании первой оценки ВВП Великобритании.

В (Lehmann, 2024) представлена база данных реального времени для региональных экономических счетов Германии, охватывающая девять макроэкономических агрегатов и 16 федеральных земель. Анализ показывает отсутствие систематических ошибок пересмотра в первых официальных оценках, однако объединение данных подавляет характеристики пересмотра отдельных регионов. Прогнозы роста валового внутреннего продукта на основе смешанной векторной авторегрессии продемонстрировали высокую точность, превосходящую эталонные модели. Исследование (Berger et al., 2023) предлагает метод прямого прогнозирования разрыва в выпуске с использованием декомпозиции Бевериджа–Нельсона и байесовской векторной авторегрессии смешанной частоты. Модель смешанной частоты обеспечивает более своевременные оценки разрыва в выпуске в США по сравнению с квартальными и другими методами. Анализ показывает, что прогнозы разрыва внутри квартала более надежны, чем прогнозы роста выпуска, а месячные индикаторы, такие как кредитный риск, потребительские настроения и уровень безработицы, предоставляют ценную информацию для окончательной оценки разрыва.

Jardet, Meunier (2022) исследуют, как высокочастотные данные могут улучшить точность прогнозирования роста мирового ВВП в условиях кризиса, например, во время пандемии COVID-19. Созданный ими набор данных включает 718 месячных и 255 недельных временных рядов, а метод FA-MIDAS с предварительным отбором переменных значительно повышает производительность модели. Результаты показывают, что модели с недельными данными превосходят другие подходы в кризисные периоды, предлагая своевременные и точные прогнозы роста мирового ВВП. В статье (Barbaglia et al., 2023) рассматриваются проблемы получения надежных наукастов во время COVID-19, когда стандартные модели устарели. Авторы предлагают два новшества: создание обширного набора данных для макроэкономического прогнозирования в Европе, включающего более тысячи временных рядов, и методологию объединения большого объема данных с классическими и современными динамическими байесовскими методами прогнозирования. Применяя эту методологию к кризису COVID-19, исследователи выявляют эффективные предикторы для прогнозирования ВВП и делают выводы для будущих кризисов.

Вопрос наукастинга макроэкономических показателей на российских данных в последние годы становится все более актуальным. Так, работа (Станкевич, 2020) посвящена изучению точности наукастов текущих темпов роста ВВП на основе оперативных данных более высокой частоты. Сравниваются различные модели, включая модификации MIDAS

с регуляризацией и метод главных компонент, а также байесовские векторные авторегрессии смешанной частоты (MFBVAR) с априорным распределением Миннесоты. Показывается, что модели на основе векторных авторегрессий превосходят другие типы моделей, анализируется динамика ошибок наукастов и приводятся оценки месячных темпов роста ВВП, полученные с помощью MFBVAR. В исследовании (Жемков, 2021) представлен метод комбинирования прогнозов для краткосрочной оценки темпов экономического роста в России, использующий около 500 моделей и учитывающий сценарные переменные и высокочастотные показатели. Такой подход показывает наилучшую точность прогнозов по сравнению с распространенными моделями за период с января 2003 по март 2020 г., что, по мнению автора, может быть полезно для разработки денежно-кредитной политики в России. В статье (Макеева, Станкевич, 2022) рассматривается оперативная оценка темпов роста ВВП России и его компонентов на основе квартальных данных с использованием различных моделей наукастинга. Сравниваются ограниченные и неограниченные MIDAS-модели, MIDAS-модель с L1-регуляризацией и MFBVAR-модель, а также классическая авторегрессия. Предложен метод оценки состояния экономики на основе данных только за первый или первые два месяца квартала, с выбором оптимальной модели по критериям средней абсолютной ошибки (MAE) и среднеквадратичной ошибки прогноза (RMSE). Фокин (2023) исследует качество наукастов и прогнозов российского ВВП с использованием модели MFBVAR, которая значительно повышает точность прогнозов для ВВП, потребления и внешней торговли по сравнению с другими моделями, особенно в кризисные периоды 2015 и 2020 гг. Работа (Зубарев, Рыбак, 2021) посвящена исследованию точности динамической факторной модели для наукастинга ВВП России. Авторы показывают, что используемая модель превосходит по точности официальные прогнозы Министерства экономического развития РФ.

Гареев и Полбин (2022) предлагают методику наукастинга и краткосрочного прогнозирования квартальных изменений ключевых макроэкономических показателей, таких как ВВП и потребление, с использованием методов машинного обучения, включая бустинг и случайный лес. Использовались данные фондового и денежного рынков, опросы и мировые цены на ресурсы. Результаты показали значительное превосходство методов машинного обучения по сравнению с наивным прогнозом и способность идентифицировать падение ВВП в условиях пандемии COVID-19 уже после десятой недели квартала. В статье (Макеева и др., 2024) исследуется наукастинг макроэкономических показателей России с использованием данных новостного фона, включая тематику и тональность новостей из крупнейших Telegram-каналов, анализируемых с помощью нейронной сети BERT. Рассматриваются модели MIDAS, динамические факторные модели и векторные авторегрессии смешанной частоты. Результаты показывают, что в кризисные периоды добавление новостных переменных повышает точность прогнозов для 45% моделей, тогда как в спокойные времена эффект менее заметен, что указывает на ограниченную роль новостного сентимента в наукастинге ВВП. В исследовании (Станкевич, 2023) рассматривается наукастинг ВВП России и его компонентов с использованием MIDAS-моделей с марковским переключением. Рассматриваются различные подходы к получению наукастов, включая взвешенные по вероятностям режимы и наиболее вероятный режим при правильном предсказании. Показано, что MIDAS-модели с марковским переключением обеспечивают максимальную точность прогнозов, и даже без правильного предсказания режима они стабильно превосходят стандартные MIDAS- и MFBVAR-модели для большинства анализируемых данных.

Вопрос пересмотра официальной статистики макроэкономических показателей на российских данных рассматривается на текущий момент только в рамках обзора вопросов прогнозирования. В работе (Турунцева, 2011) представлен обзор основных моделей, применяемых в России для прогнозирования макроэкономических показателей. Подчеркивается, что эконометрические модели являются основным методом прогнозирования, обеспечивающим надежные оценки ключевых макроэкономических индикаторов. Астафьева, Турунцева (2021) предлагают методику сбора данных по плановым и внеплановым переоценкам номинального объема ВВП и индекса физического объема ВВП с I квартала 1994 г. по I квартал 2020 г., публикуемых в сборнике «Краткосрочные экономические показатели Российской Федерации». Авторы сформировали винтажи этих показателей в годовом и квартальном разрезах, предложили варианты русскоязычных терминов, а также провели анализ статистических свойств пересмотров показателей ВВП. В частности, исследованы систематические смещения в переоценках, изменения динамических свойств и наличие шумовых и новостных компонентов в показателях валового внутреннего продукта.

В настоящем исследовании поднимается вопрос влияния версии ревизии официальной статистики в зависимости от доступности данных внутри квартала на точность моделей наукастинга, применяемых к макроэкономическим показателям России. Насколько известно автору, данная работа является первой в своем роде с использованием российских данных, что дополняет отечественные исследования вопросов наукастинга и краткосрочного прогнозирования.

2. Методы

Целью данного исследования является сравнение точности вневыборочных наукастов макроэкономических показателей России в зависимости от версии ревизии официальной статистики. В данной работе применяются модели, показавшие наилучшую точность при наукастинге макроэкономических показателей, включающие как эконометрические подходы, так и техники, близкие с методами машинного обучения.

Одним из наиболее известных и доступных методов для краткосрочной оценки макроэкономических показателей с применением данных разной частоты является MIDAS-модель. Этот тип модели относится к эконометрическим методам прогнозирования, и формально его можно представить следующим образом:

$$y_t = \sum_{j=1}^p \alpha_j y_{t-j} + \sum_{i=0}^k \sum_{j=0}^{m_i} \beta_j^{(i)} x_{tm_i-j}^{(i)} + u_t,$$

где y_t — это прогнозируемая переменная, представляющая собой ряд с низкой частотой (в данном случае — квартальные темпы роста по сравнению с аналогичным периодом предыдущего года), $x_t^{(i)}$ — регрессоры, которые являются рядами с высокой частотой (месячные темпы роста по сравнению с аналогичным периодом прошлого года), m_i — число наблюдений переменной $x_t^{(i)}$ в одном временном интервале для прогнозируемой переменной y_t .

Следует подчеркнуть, что для упомянутой выше MIDAS-модели характерно отсутствие ограничений на параметры, что и стало причиной ее названия — неограниченная MIDAS-модель (U-MIDAS-модель). На практике более широкое применение нашли MIDAS-модели,

в которых параметры $\beta_j^{(i)}$ ограничены различными способами. В общем виде такие ограничения можно выразить следующим образом:

$$\beta_j^{(i)} = g^{(i)}(j, \lambda),$$

где вид функции $g^{(i)}$ зависит от типа ограничений, применяемых к U-MIDAS-модели. Разные ограничения на параметры модели служат методами агрегации данных более высокой частоты в данные низкой частоты и направлены на повышение точности прогнозов. В данном исследовании ограничения, накладываемые на U-MIDAS-модель, будут агрегировать месячные данные в квартальные.

Рассмотрим различные MIDAS-модели с ограничениями (MIDAS-R-модели), которые продемонстрировали наивысшую эффективность в краткосрочном прогнозировании различных макроэкономических индикаторов:

- MIDAS-R-модель с неэкспоненциальными полиномами Алмона (Clements, Galvão, 2009)

$$\beta_j = \frac{\exp\{\lambda_1(j+1) + \dots + \lambda_r(j+1)^r\}}{\sum_{s=0}^d \exp\{\lambda_1(s+1) + \dots + \lambda_r(s+1)^r\}};$$

- MIDAS-R-модель с экспоненциальными полиномами Алмона (MIDAS-R-almonp-модель)

$$\beta_j = \sum_{s=0}^d \lambda_s j^s;$$

- MIDAS-R-модель с ограничением на основе распределения Гомпертца (MIDAS-R-gompertzp-модель)

$$\beta_j = \frac{\exp\{\lambda_2 j - \lambda_1 e^{\lambda_2 j}\}}{\sum_{s=0}^d \exp\{\lambda_2 s - \lambda_1 e^{\lambda_2 s}\}}.$$

Необходимо отметить, почему MIDAS-R-модели в различных вариантах демонстрируют большую точность по сравнению с классической U-MIDAS-моделью. Ограничения, применяемые к параметру, направлены на уменьшение размерности данных, что, с одной стороны, помогает избежать переобучения модели (термин машинного обучения), а с другой стороны, позволяет сохранить вариативность модели для дальнейшей адаптации входных наборов данных. Для создания указанных моделей и оценки их эффективности применяется пакет *midasr* (Ghysels et al., 2016) в языке программирования R.

Второй тип моделей, применяемых для оперативной оценки макроэкономических показателей, представляет собой байесовскую векторную авторегрессию смешанной частоты (Mixed-Frequency Bayesian VAR, MFBVAR). В данной работе используется вариант байесовской векторной авторегрессии с априорным распределением Миннесоты. Эту модель можно рассматривать как усовершенствованную версию классической векторной авторегрессии, адаптированную для работы с данными различной периодичности (в рамках данного исследования — месячные и квартальные данные). Механизм работы априорного распределения подробно описан в работах (Litterman, 1979, 1986).

MFBVAR-модель позволяет не только включать широкий спектр объясняющих переменных, но и использовать лаги этих переменных при недостатке входных данных. Это становится

возможным благодаря автоматическому исключению из модели тех переменных, которые не влияют на точность краткосрочного прогноза. В данной работе была выбрана MFBVAR-модель с априорным распределением Миннесоты, т.к. использование других типов распределений требует более сложных априорных предположений (например, спецификации ковариационных структур или иерархических параметров), что увеличивает риск переобучения и снижает точность прогнозов. Для учета полного набора объясняющих переменных в одной модели они обрабатываются не в исходном виде, а как первые три главные компоненты.

Модификациями моделей с возможностью исключения переменных являются MIDAS-модели с L1 (LASSO) и L2 (Ridge) регуляризацией

$$y_t = \beta^0 + \sum_{i=1}^k \sum_{j=0}^{m_i} \beta_j^{(i)} x_{t, m_i-j}^{(i)} + u_t$$

с целевыми функциями:

$$\sum_{t=1}^T (y_t - \hat{y}_t)^2 + \lambda \sum_{i=1}^k \sum_{j=0}^{m_i} |\beta_j^{(i)}| \rightarrow \min \text{ для LASSO-U-MIDAS модели (MIDAS_L1-модель),}$$

$$\sum_{t=1}^T (y_t - \hat{y}_t)^2 + \lambda \sum_{i=1}^k \sum_{j=0}^{m_i} (\beta_j^{(i)})^2 \rightarrow \min \text{ для Ridge-U-MIDAS модели (MIDAS_L2-модель),}$$

где λ — параметр, который служит штрафом за добавление большого числа переменных и определяется с помощью кросс-валидации. В данном случае применяется значение $\lambda_{\min}$, соответствующее минимальному значению RSS при кросс-валидации.

Третьей и последней линейной моделью с регуляризацией в этом исследовании является модель с методом главных компонент MIDAS_PC. В ней вместо исходного набора объясняющих переменных используются первые пять главных компонент, что значительно уменьшает размерность задачи при минимальных потерях информации из исходных данных.

Линейные модели с регуляризацией основаны на классической U-MIDAS-модели, которая накладывает ограничения, позволяя включать лишь одну объясняющую переменную. При добавлении второй переменной число оцениваемых коэффициентов значительно увеличивается. Например, если в моделях со смешанной частотой используются месячные данные в качестве объясняющих и квартальные в качестве прогнозируемой переменной, то добавление еще одной объясняющей переменной требует включения трех дополнительных регрессоров. Линейные модели с регуляризацией помогают смягчить эту проблему, позволяя использовать больше переменных для получения прогнозов.

Динамические факторные модели (DFM-модели) предполагают, что динамику макроэкономической переменной можно объяснить с помощью ограниченного числа ненаблюдаемых факторов. В качестве таких факторов могут использоваться, например, оперативно публикуемые официальные статистические данные. Одним из преимуществ DFM-моделей является решение проблемы «проклятия размерности», т.к. в модель входит небольшое количество данных. В данной работе рассматривается DFM, описанная в (Doz et al., 2011; Giannone et al., 2008).

3. Данные и оценка точности моделей

В настоящей работе модели наукастинга строятся для квартальных данных ВВП России и его компонентов по использованию, которые публикует Федеральная служба государственной статистики. Оценка моделей проводилась с первого квартала 2014 г. по третий квартал 2023 г. Для формирования полных временных рядов набор данных с базовыми ценами 2016 г. был дополнен информацией за более ранние периоды с другими базовыми ценами после необходимого пересчета. Зависимые переменные имеют квартальную частоту, а объясняющие — месячную. В качестве объясняющих переменных были выбраны ключевые индикаторы, отражающие экономическую ситуацию в стране, а в качестве регрессоров добавлены цены на основные энергоносители. Полный список, описание переменных, календарь выхода и источник данных представлены в табл. П1 и П2 в Приложении. Все переменные были переведены в темпы роста по сравнению с аналогичным периодом предыдущего года для устранения сезонной компоненты и обеспечения стационарности показателей.

В качестве основного показателя для оценки качества прогнозов предлагается использовать среднюю абсолютную ошибку (MAE):

$$MAE = \frac{1}{T} \sum_{t=1}^T |Y_t - \hat{Y}_t|,$$

где Y_t — фактическая наблюдаемая зависимая переменная (в данном случае — прирост по сравнению с аналогичным периодом предыдущего года), $\hat{Y}_t$ — предсказанное значение, полученное с помощью модели, T — число периодов, по которым проводится оценка точности модели.

В рамках данной работы тестировалась точность вневыборочных наукастов. Сравнение моделей по показателю RMSE (средняя квадратичная ошибка прогноза) не приводится, поскольку результаты совпали с аналогичными по показателю MAE.

Оценка точности моделей берется по последним 12 точкам с учетом следующих особенностей.

1. Модели независимо оцениваются на первой и финальной ревизиях официальной статистики. Первой ревизией в данной работе называется первое опубликованное значение рассматриваемого показателя, а финальной — последнее доступное на момент проведения исследования.

2. Результат наукастинга переменных сравнивается со всеми доступными ревизиями официальной статистики для расчета точности используемых моделей и оценки влияния версий ревизии на точность моделей наукастинга.

3. Исследуется, насколько полезны месячные данные в пределах прогнозируемого периода. Создаются модели прогнозирования с использованием данных только за первый месяц, за первые два месяца или за весь квартал.

4. В качестве бенчмарка используется авторегрессия первого порядка.

4. Результаты

Пересмотр версий официальной статистики может происходить как один раз, когда для потребителей данных становится доступно только первое и финальное значения, так и содержать большее число итераций. Например, на рис. 1 представлена статистика по пересмотрам квартальных версий официальной статистики по ВВП России и его компонентам по использованию с начала 2018 г. Данные показывают, что больше всего пересматривается основная переменная — ВВП России (до четырех версий официальной статистики за рассматриваемый период, до трех версий официальной статистики на периоде тестирования моделей), остальные переменные, в основном, пересматриваются не более двух раз.

Если же рассматривать годовые версии официальной статистики (рис. 2), то с 2018 г. данные пересматривались один или два раза, формируя два или три значения интересующих показателей соответственно.

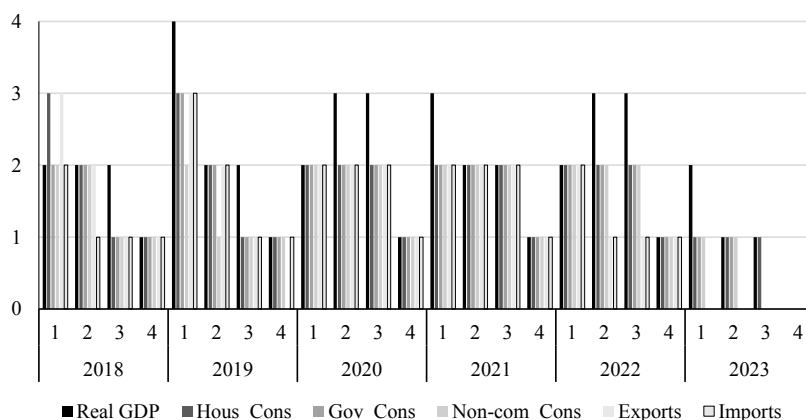


Рис. 1. Статистика по пересмотрам квартальных версий официальной статистики по ВВП России и его компонентам по использованию

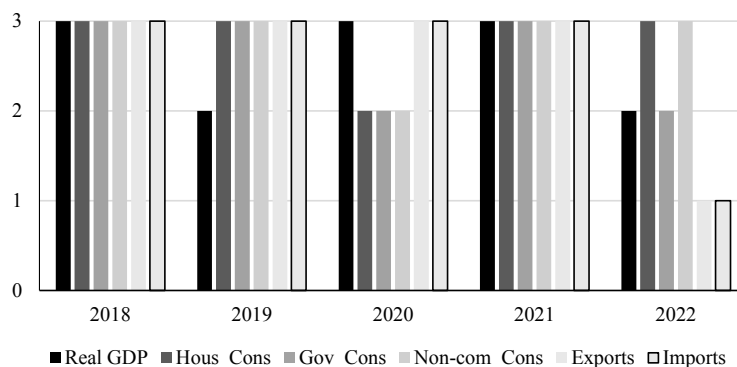


Рис. 2. Статистика по пересмотрам годовых версий официальной статистики по ВВП России и его компонентам по использованию

В таблицах 1–6 представлены средние абсолютные ошибки наукастов для всех моделей и переменных. Если за рассматриваемый период (12 точек) было три версии официальной статистики, то сравниваются первая, вторая и финальная ревизии. Если же было всего две версии официальной статистики, то сравнивается точность моделей наукастинга для первой и финальной ревизий. Для всех модификаций MIDAS-моделей указан наилучший результат среди всех объясняющих переменных, а в скобках указывается переменная, показавшая наилучшую точность. Используются следующие обозначения для моделей:

- U-MIDAS — неограниченная MIDAS-модель;
- MIDAS-R-nealmon — MIDAS-модель с неэкспоненциальными полиномами Алмона;
- MIDAS-R-almonp — MIDAS-модель с экспоненциальными полиномами Алмона;
- MIDAS-R-gompertzp — MIDAS-модель с ограничением на основе распределения Гомпертца;
- MFBVAR — байесовская векторная авторегрессия смешанной частоты;
- MIDAS_L1 — MIDAS-модель с полным набором объясняющих переменных и LASSO-регуляризацией;
- MIDAS_L2 — MIDAS-модель с полным набором объясняющих переменных и Ridge-регуляризацией;
- MIDAS_PC — MIDAS-модель с полным набором объясняющих переменных и регуляризацией на основе метода главных компонент;
- DFM — динамическая факторная модель;
- AR(1) — авторегрессия первого порядка.

Анализ моделей наукастинга для годовых темпов роста ВВП России по использованию (переменная *GDP*) показывает, что наибольшую точность обнаруживает MIDAS-модель с ограничением на основе распределения Гомпертца, а наибольшая точность достигается при использовании данных за полный квартал вне зависимости от того, на каких данных

Таблица 1. MAE наукастов для годовых темпов роста ВВП России по использованию

На каких данных оценивалась модель	Полнота данных	На каких данных рассчитывалась точность					
		Первая ревизия		Вторая ревизия		Финальное значение	
		MAE	Модель	MAE	Модель	MAE	Модель
Первая ревизия	3 месяца	0.61	MIDAS-R-gompertzp (Basic_Output)	0.60	MIDAS-R-gompertzp (Basic_Output)	0.62	MIDAS-R-gompertzp (Basic_Output)
	2 месяца	0.63	MIDAS-R-gompertzp (Basic_Output)	0.68	MIDAS-R-gompertzp (Basic_Output)	0.73	MIDAS-R-gompertzp (Basic_Output)
	1 месяц	0.83	MIDAS-R-gompertzp (Basic_Output)	0.93	MIDAS-R-gompertzp (Basic_Output)	0.98	MIDAS-R-gompertzp (Basic_Output)
Финальное значение	3 месяца	0.65	MIDAS-R-gompertzp (Basic_Output)p	0.65	MIDAS-R-gompertzp (Basic_Output)p	0.67	MIDAS-R-gompertzp (Basic_Output)p
	2 месяца	0.63	MIDAS-R-almonp (Basic_Output)	0.69	MIDAS-R-almonp (Basic_Output)	0.74	MIDAS-R-almonp (Basic_Output)
	1 месяц	0.84	MIDAS-R-almonp (Basic_Output)	0.92	MIDAS-R-almonp (Basic_Output)	0.97	MIDAS-R-almonp (Basic_Output)

Таблица 2. MAE наукастов для годовых темпов роста расходов на конечное потребление домашних хозяйств

На каких данных оценивалась модель	Полнота данных	На каких данных рассчитывалась точность			
		Первая ревизия		Финальное значение	
		MAE	Модель	MAE	Модель
Первая ревизия	3 месяца	3.20	MIDAS-R-gompertzp (Basic_Output)	3.01	MIDAS-R-gompertzp (Basic_Output)
	2 месяца	3.32	MIDAS-R-gompertzp (Manufact)	3.18	MIDAS-R-gompertzp (Manufact)
	1 месяц	3.26	MIDAS-R-gompertzp (Manufact)	3.12	MIDAS-R-gompertzp (Manufact)
Финальное значение	3 месяца	3.24	MIDAS-R-gompertzp (Manufact)	3.11	MIDAS-R-gompertzp (Manufact)
	2 месяца	3.19	MIDAS-R-gompertzp (Manufact)	3.06	MIDAS-R-gompertzp (Manufact)
	1 месяц	3.16	MIDAS-R-gompertzp (Manufact)	3.03	MIDAS-R-gompertzp (Manufact)

Таблица 3. MAE наукастов для годовых темпов роста расходов на конечное потребление государственного управления

На каких данных оценивалась модель	Полнота данных	На каких данных рассчитывалась точность			
		Первая ревизия		Финальное значение	
		MAE	Модель	MAE	Модель
Первая ревизия	3 месяца	2.32	MIDAS_L2	2.66	MIDAS-R-gompertzp (Wage)
	2 месяца	2.29	MIDAS_L1	2.45	MIDAS_L1
	1 месяц	2.38	MIDAS_L2	2.53	MIDAS-R-gompertzp (Wage)
Финальное значение	3 месяца	2.14	MIDAS_L2	2.41	MIDAS-R-nealmon (Unempl)
	2 месяца	2.15	MIDAS_L2	2.19	MIDAS-R-nealmon (Wage)
	1 месяц	2.19	MIDAS_L2	2.25	MIDAS-R-almonp (Wage)

оценивалась модель (MAE варьируется от 0.60 до 0.67). Если же говорить о точности моделей в зависимости от версии ревизии официальной статистики, то наибольшая точность достигается при наукастинге ее первой версии.

Для годовых темпов роста расходов на конечное потребление домашних хозяйств (переменная *House_Cons*) наилучшая точность также достигается на MIDAS-модели с ограничением на основе распределения Гомпертца. При использовании модели, оцененной на первой ревизии официальной статистики, лучшую точность показывает набор данных за все три месяца рассматриваемого квартала (MAE составила 3.20 при сравнении с первой ревизией и 3.01 при сравнении с остальными), а при оценке модели на финальной ревизии лучшая точность наблюдается при наличии данных только за первый месяц рассматриваемого квартала (MAE составила 3.16). При сравнении точности между второй, третьей и финальной ревизиями разницы не обнаружено вне зависимости от того, на каких данных обучалась модель.

При наукастинге годовых темпов роста расходов на конечное потребление государственного управления (переменная *Gov_Cons*) наилучшие результаты достигаются с использованием модели с регуляризацией L1, особенно при оценке на первой ревизии официальной статистики, где MAE составляет 2.29 для первой ревизии при использовании данных за два

Таблица 4. MAE наукастов для годовых темпов роста расходов на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства

На каких данных оценивалась модель	Полнота данных	На каких данных рассчитывалась точность			
		Первая ревизия		Финальное значение	
		MAE	Модель	MAE	Модель
Первая ревизия	3 месяца	1.72	MIDAS-R-gompertzp (House)	2.75	DFM
	2 месяца	1.78	MIDAS-R-nealmon (Unempl)	2.84	MIDAS-R-gompertzp (Catering)
	1 месяц	1.78	MIDAS-R-nealmon (Unempl)	2.85	MIDAS-R-gompertzp (Catering)
Финальное значение	3 месяца	1.53	MIDAS-R-almonp (Catering)	1.63	MFBVAR
	2 месяца	1.62	MIDAS-R-gompertzp (Retail_Food)	1.75	AR(1)
	1 месяц	1.59	MIDAS-R-almonp (Catering)	1.75	AR(1)

месяца. При оценке на финальной ревизии модель L2 показывает наименьшую ошибку (2.14) также при использовании трехмесячных данных. Используемые модели в целом показывают лучшую точность при наукастинге первой ревизии официальной статистики.

Для годовых темпов роста расходов на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства (переменная *Non-com_Cons*), наилучшие результаты показывают модели MIDAS-R-gompertzp и MIDAS-R-almonp, особенно при оценке на финальной ревизии официальной статистики. Наиболее низкое значение MAE (1.53) достигается с использованием модели MIDAS-R-almonp при наличии данных за три месяца. При оценке с двумя месяцами данных модель MIDAS-R-gompertzp также демонстрирует хорошую сравнительную точность с MAE 1.62. Важно отметить, что все модели показывают лучшие результаты при использовании трехмесячных данных, что подчеркивает важность полноты информации для повышения точности прогнозов. Более того, рассматриваемые модели наукастинга лучше прогнозируют первую ревизию официальной статистики по сравнению с финальной.

Самый точный краткосрочный прогноз годовых темпов роста экспорта (переменная *Exports*) достигается с использованием модели MFBVAR при оценке на первой ревизии официальной статистики, где MAE составляет 7.25 при использовании данных за три месяца. При оценке с двумя месяцами данных MFBVAR-модель демонстрирует значительное улучшение точности при оценке как на первой, так и на финальной ревизии статистики, что подчеркивает ее эффективность при меньшей полноте данных для рассматриваемой

Таблица 5. MAE наукастов для годовых темпов роста экспорта

На каких данных оценивалась модель	Полнота данных	На каких данных рассчитывалась точность			
		Первая ревизия		Финальное значение	
		MAE	Модель	MAE	Модель
Первая ревизия	3 месяца	7.46	MIDAS-R-gompertzp (Mining)	7.25	MFBVAR
	2 месяца	6.16	MFBVAR	5.88	MFBVAR
	1 месяц	5.93	MFBVAR	5.57	MFBVAR
Финальное значение	3 месяца	7.60	MIDAS-R-gompertzp (Mining)	7.52	MFBVAR
	2 месяца	6.24	MFBVAR	5.84	MFBVAR
	1 месяц	6.28	MFBVAR	5.93	MFBVAR

Таблица 6. MAE наукастов для годовых темпов роста импорта

На каких данных оценивалась модель	Полнота данных	На каких данных рассчитывалась точность			
		Первая ревизия		Финальное значение	
		MAE	Модель	MAE	Модель
Первая ревизия	3 месяца	5.78	MIDAS-R-gompertzp (Basic_Output)	5.97	MIDAS-R-gompertzp (Basic_Output)
	2 месяца	7.32	MIDAS-R-gompertzp (Manufact)	7.62	MIDAS-R-gompertzp (Manufact)
	1 месяц	8.24	MIDAS-R-gompertzp (Wholesale)	8.14	MIDAS-R-gompertzp (Wholesale)
Финальное значение	3 месяца	5.71	MIDAS-R-nealmon (Prom)	5.99	MIDAS-R-nealmon (Prom)
	2 месяца	7.23	MIDAS-R-almonp (Manufact)	7.53	MIDAS-R-almonp (Manufact)
	1 месяц	8.10	MIDAS-R-gompertzp (Wholesale)	8.08	MIDAS-R-gompertzp (Wholesale)

переменной. Следует отметить, что MFBVAR-модель демонстрирует стабильно лучшую точность для переменной *Exports* среди рассматриваемых моделей наукастинга.

При наукастинге годовых темпов роста импорта (переменная *Imports*) наилучшие результаты достигаются с использованием MIDAS-модели с ограничениями, особенно при оценке на финальной ревизии официальной статистики, где MAE составляет 5.71 при использовании данных за три месяца. В первой ревизии модель MIDAS-R-gompertzp показывает MAE 5.78, что также является хорошим результатом, однако в финальной ревизии ее точность снижается до 8.10 при использовании одного месяца данных. При оценке с двумя месяцами модель MIDAS-R-almonp демонстрирует MAE 7.23, что лучше, чем у других моделей. Для всех рассматриваемых вариантов наличие данных за полный квартал значительно улучшает точность наукаста.

На основе полученных расчетов можно сделать следующие выводы (табл. 7, 8):

1. Для получения лучшего наукаста первой ревизии необходимо использовать модель, оцененную по первой ревизии официальной статистики с данными за полный квартал для годовых темпов роста ВВП России по использованию, а для годовых темпов экспорта — с данными только за первый месяц рассматриваемого квартала. Для остальных же переменных для получения наиболее точного наукаста первой ревизии официальной статистики необходимо использовать модель, оцененную на финальных значениях. При этом для годовых темпов роста расходов на конечное потребление домашних хозяйств лучший наукаст получается при использовании данных за первый месяц, для остальных — за три месяца рассматриваемого квартала.

2. Для получения наиболее точного наукаста финального значения официальной статистики для годовых темпов роста ВВП России по использованию, расходов на конечное потребление домашних хозяйств, экспорта и импорта необходимо использовать модель, оцененную по первым ревизиям официальной статистики. При этом для годовых темпов роста экспорта лучше использовать данные за первый месяц рассматриваемого квартала, а для других переменных — за полный квартал. При наукастинге годовых темпов роста расходов на конечное потребление государственного управления и расходов на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства, лучше использовать модель, оцененную на финальных значениях статистики, для первой переменной использовать данные за два месяца, а для второй — за три месяца рассматриваемого квартала.

3. С точки зрения используемых моделей при наукастинге первой версии официальной статистики достаточно точной оказывается MIDAS-модель с ограничениями. При этом

Таблица 7. Какие модели и данные лучше использовать для получения наиболее точного наукаста первой ревизии официальной статистики по ВВП России и его компонентам по использованию

	Переменная					
	<i>GDP</i>	<i>Hous Cons</i>	<i>Gov Cons</i>	<i>Non-com Cons</i>	<i>Exports</i>	<i>Imports</i>
Модель	MIDAS-R-gompertzp (Basic_Output)	MIDAS-R-gompertzp (Manufact)	MIDAS_L2	MIDAS-R-almonp (Catering)	MFBVAR	MIDAS-R-nealmon (Prom)
На каких данных оценивалась модель?	Первая ревизия	Финальное значение	Финальное значение	Финальное значение	Первая ревизия	Финальное значение
Полнота данных	3 месяца	1 месяц	3 месяца	3 месяца	1 месяц	3 месяца

для годовых темпов роста ВВП России по использованию и расходов на конечное потребление домашних хозяйств лучше работают ограничения на основе распределения Гомпертца, а для годовых темпов роста расходов на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства, и импорта — с экспоненциальными и неэкспоненциальными полиномами Алмона соответственно. Для годовых темпов роста расходов на конечное потребление государственного управления при наукастинге первой ревизии наиболее точной получилась L2-модель, а для годовых темпов роста экспорта — MFBVAR-модель.

4. Для наукастинга финального значения официальной статистики хорошую точность также показывают MIDAS-модели с ограничениями. Для годовых темпов роста ВВП России по использованию, расходов на конечное потребление домашних хозяйств и импорта лучше работает ограничение на основе распределения Гомпертца, а для годовых темпов роста расходов на конечное потребление государственного управления — ограничение на основе неэкспоненциальных полиномов Алмона. Для двух оставшихся переменных (годовые темпы роста расходов на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства, и годовые темпы роста экспорта) при наукастинге финального значения лучше всего работает MFBVAR-модель.

Таблица 8. Какие модели и данные лучше использовать для получения наиболее точного наукаста финального значения официальной статистики по ВВП России и его компонентам по использованию

	Переменная					
	<i>GDP</i>	<i>Hous Cons</i>	<i>Gov Cons</i>	<i>Non-com Cons</i>	<i>Exports</i>	<i>Imports</i>
Модель	MIDAS-R-gompertzp (Basic_Output)	MIDAS-R-gompertzp (Basic_Output)	MIDAS-R-nealmon (Wage)	MFBVAR	MFBVAR	MIDAS-R-gompertzp (Basic_Output)
На каких данных оценивали модель?	Первая ревизия	Первая ревизия	Финальное значение	Финальное значение	Первая ревизия	Первая ревизия
Полнота данных	3 месяца	3 месяца	2 месяца	3 месяца	1 месяц	3 месяца

5. Переменная, отражающая годовые темпы роста фактического объема выпуска по базовым видам экономической деятельности, показывает наилучшую точность при использовании ее в качестве объясняющей переменной в MIDAS-моделях.

6. Несмотря на то что первая ревизия официальной статистики многократно пересматривается Росстатом, модель наукастинга для ВВП России, оцененная на первой ревизии, оказывается более точной при любой полноте данных внутри квартала.

7. Модели наукастинга для ВВП России лучше всего попадают прогнозом в первую, самую «шумную» ревизию.

8. Прогнозы первой ревизии ВВП на данных за первый месяц значительно точнее финальных наукастов, в то время как для данных за полный квартал точность всех ревизий оказывается схожей. Поскольку первые ревизии в большей степени опираются на данные, полученные в начале отчетного периода, это подчеркивает важность оперативной информации для первоначальных оценок.

5. Заключение

В работе представлены результаты анализа точности моделей наукастинга для ВВП России и его компонентов по использованию за период с 1 квартала 2014 г. по 3 квартал 2023 г. по последним 12 точкам на основе средней абсолютной ошибки прогноза. В исследовании рассмотрена проблема многократного пересмотра официальной статистики, например, квартальный ВВП планово пересматривается Росстатом до четырех раз до получения финального значения. При этом при наукастинге ВВП России и его компонентов по использованию модели оцениваются как на первой, так и на финальной ревизиях официальной статистики с учетом доступности данных внутри прогнозируемого квартала. Результаты работы демонстрируют, что точность моделей наукастинга существенно зависит от выбора версии ревизии официальной статистики, а также от полноты данных, доступных внутри прогнозируемого квартала. Для рассматриваемых макроэкономических показателей нет универсального ответа, какую модель и с какими данными использовать, однако в рамках исследования отмечается несколько закономерностей. Так, переменная, отражающая годовые темпы роста фактического объема выпуска по базовым видам экономической деятельности, показывает наилучшую точность при использовании ее в качестве объясняющей переменной в MIDAS-моделях.

Полученные результаты позволяют сделать несколько выводов о наукастинге ВВП России. Несмотря на то что первая ревизия официальной статистики подвергается многократным пересмотрам со стороны Росстата, модель наукастинга для ВВП России, основанная на первой ревизии, демонстрирует большую точность при любой полноте данных в рамках квартала. Более того, прогнозы для ВВП России наиболее точно соответствуют первой, самой «шумной» ревизии. Прогнозы по первой ревизии ВВП на основе данных за первый месяц оказываются значительно более точными, чем финальные прогнозы, в то время как при использовании данных за весь квартал точность всех ревизий оказывается схожей. Это свидетельствует о том, что первые ревизии в большей степени зависят от данных, полученных в начале отчетного периода, и подчеркивает значимость оперативной информации для первоначальных оценок.

Полученные результаты могут быть использованы для разработки более эффективных стратегий прогнозирования макроэкономических показателей, что особенно важно для принятия решений в условиях неопределенности и изменчивости экономической среды.

В дальнейших исследованиях целесообразно рассмотреть вопросы о новых методах обработки данных смешанной частоты и интеграцию методов машинного обучения в традиционные эконометрические модели для повышения их прогностической способности.

Список литературы

- Астафьева Е. В., Турунцева М. Ю. (2021). Пересмотры ВВП: данные и оценка статистических свойств. *Экономический журнал ВШЭ*, 25 (1), 65–101. DOI: 10.17323/1813-8691-2021-25-1-65-101.
- Гареев М. Ю., Полбин А. В. (2022). Наукастинг: оценка изменения ключевых макроэкономических показателей с использованием методов машинного обучения. *Вопросы экономики*, 8, 133–157. DOI: 10.32609/0042-8736-2022-8-133-157.
- Жемков М. И. (2021). Краткосрочная оценка ВВП России методом комбинирования прогнозов. *Серия докладов об экономических исследованиях, Банк России, Февраль 2021*.
- Зубарев А. В., Рыбак К. С. (2021). Наукастинг ВВП: динамическая факторная модель и официальные прогнозы. *Экономическое развитие России*, 28 (12), 34–40.
- Макеева Н. М., Станкевич И. П. (2022). Наукастинг элементов использования ВВП России. *Экономический журнал ВШЭ*, 26 (4), 598–622. DOI: 10.17323/1813-8691-2022-26-4-598-622.
- Макеева Н. М., Станкевич И. П., Любайкин Н. С. (2024). Наукастинг макроэкономических показателей экономики России в условиях неопределенности: помогает ли учет новостного фона? *Вопросы экономики*, 3, 120–142. DOI: 10.32609/0042-8736-2024-3-120-142.
- Станкевич И. П. (2020). Сравнение методов наукастинга макроэкономических индикаторов на примере российского ВВП. *Прикладная эконометрика*, 59, 113–127. DOI: 10.22394/1993-7601-2020-59-113-127.
- Станкевич И. П. (2023). Применение MIDAS-моделей с марковским переключением для наукастинга ВВП и его компонентов. *Прикладная эконометрика*, 70, 122–143. DOI: 10.22394/1993-7601-2023-70-122-143.
- Турунцева М. Ю. (2011). Прогнозирование в России: обзор основных моделей. *Экономическая политика*, 1, 193–202.
- Фокин Н. Д. (2023). Наукастинг и прогнозирование основных российских макроэкономических показателей с помощью MFBVAR-модели. *Экономическая политика*, 18 (3), 110–135. DOI: 10.18288/1994-5124-2023-3-110-135.
- Aastveit K. A., Gerdrup K. R., Jore A. S., Thorsrud L. A. (2014). Nowcasting GDP in real time: A density combination approach. *Journal of Business and Economic Statistics*, 32 (1), 48–68. DOI: 10.1080/07350015.2013.844155.
- Anesti N., Galvão A. B., Miranda-Agrippino S. (2022). Uncertain Kingdom: Nowcasting GDP and its revisions. *Journal of Applied Economics*, 37 (1), 42–62. DOI: 10.1002/jae.2845.
- Auroba S. B. (2008). Data revisions are not well-behaved. *Journal of Money Credit and Banking*, 40, 319–340. DOI: 10.1111/j.1538-4616.2008.00115.x.
- Barbaglia L., Frattarolo L., Onorante L., Pericoli F. M., Ratto M., Pezzoli L. T. (2023). Testing big data in a big crisis: Nowcasting under Covid-19. *International Journal of Forecasting*, 39 (4), 1548–1563. DOI: 10.1016/j.ijforecast.2022.10.005.

- Berger T., Morley J., Wong B. (2023). Nowcasting the output gap. *Journal of Econometrics*, 232 (1), 18–34. DOI: 10.1016/j.jeconom.2020.08.011.
- Carriero A., Clements M. P., Galvão A. B. (2015). Forecasting with Bayesian multivariate vintage-based VARs. *International Journal of Forecasting*, 35, 757–768. DOI: 10.1016/j.ijforecast.2014.05.007.
- Cepni O., Guney I., Swanson N. (2020) Forecasting and nowcasting emerging market GDP growth rates: The role of latent global economic policy uncertainty and macroeconomic data surprise factors. *Journal of Forecasting*, 39 (1), 18–36. DOI: 10.1002/for.2602.
- Chikamatsu K., Hirakata N., Kido Y., Otaka K. (2018). Nowcasting Japanese GDPs. *Bank of Japan Working Paper Series 18-E-18*. Bank of Japan.
- Clarida R., Kitchen J., Monaco R. (2002). Real-time forecasting in practice: A description of the U.S. Treasury's real-time GDP forecasting system. *MPRA Paper 21068, University Library of Munich, Germany*.
- Clements M. P., Galvão A. B. (2009). Forecasting US output growth using leading indicators: An appraisal using MIDAS models. *Journal of Applied Econometrics*, 24 (7), 1187–1206. DOI: 10.1002/jae.1075.
- Clements M. P., Galvão A. B. (2012). Improving real-time estimates of output gaps and inflation trends with multiple-vintage VAR models. *Journal of Business and Economic Statistics*, 30 (4), 554–562. DOI: 10.2307/41739785.
- Clements M. P., Galvão A. B. (2017). Predicting early data revisions to U.S. output growth and the effects of releases on equity markets. *Journal of Business and Economic Statistics*, 35 (3), 591–604. DOI: 10.1080/07350015.2015.1076726.
- Croushore D. (2008). Revisions to PCE inflation measures: Implications for monetary policy. *Working paper no. 08–8*. Federal Reserve Bank of Philadelphia.
- Dahlhaus T., Guénette J., Vasishtha G. (2017). Nowcasting BRIC+M in real time. *International Journal of Forecasting*, 33 (4), 915–935. DOI: 10.1016/j.ijforecast.2017.05.002.
- Doz C., Giannone D., Reichlin L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, 164 (1), 188–205. DOI: 10.1016/j.jeconom.2011.02.012.
- Evans M. D. D. (2005). Where are we now? *International Journal of Central Banking*, 1, 127–175.
- Galvão A. B. (2017). Data revisions and DSGE models. *Journal of Econometrics*, 196 (1), 215–232. DOI: 10.1016/j.jeconom.2016.09.006.
- Galvão A. B., Mitchell J. (2019). Measuring data uncertainty: An application using the Bank of England's "fan charts" for historical GDP growth. *EMF Working Paper Series*, 24.
- Giannone D., Reichlin L., Small D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55 (4), 665–676.
- Giovannelli A., Proietti T., Citton A., Ricchi O., Tegami C., Tinti C. (2020). Nowcasting GDP and its components in a data-rich environment: The merits of the indirect approach. *CEIS Research Paper*, 489. DOI: 10.2139/ssrn.3614110.
- Ghysels E., Kvedaras V., Zemlys V. (2016). Mixed frequency data sampling regression models: The R package midasr. *Journal of Statistical Software, Foundation for Open Access Statistics*, 72, 1–35. DOI: <http://hdl.handle.net/10.18637/jss.v072.i04>.

- Hayashi F., Tachi Y. (2021). Generalized nowcast revision analysis. DOI: 10.2139/ssrn.3904418.
- Hayashi F., Tachi Y. (2023). Nowcasting Japan's GDP. *Empirical Economics*, 64 (4), 1699–1735. DOI: 10.1007/s00181-022-02301-w.
- Howrey E. P. (1978). The use of preliminary data in economic forecasting. *The Review of Economics and Statistics*, 60 (2), 193–200.
- Howrey E. P. (1984). Data revisions, reconstruction and prediction: An application to inventory investment. *The Review of Economics and Statistics*, 66 (3), 386–393.
- Jacobs J. P. A. M., van Norden S. (2011). Modeling data revisions: Measurement error and dynamics of 'true' values. *Journal of Econometrics*, 161 (2), 101–109. DOI: 10.1016/j.jeconom.2010.04.010.
- Jardet C., Meunier B. (2022). Nowcasting world GDP growth with high-frequency data. *Journal of Forecasting*, 41 (6), 1181–1200. DOI: 10.1002/for.2858.
- Kishor N. K., Koenig E. F. (2012). VAR estimation and forecasting when data are subject to revision. *Journal of Business and Economic Statistics*, 30 (2), 181–190. DOI: 10.1198/jbes.2010.08169.
- Kočenda E., Poghosyan K. (2020). Nowcasting real GDP growth: Comparison between old and new EU countries. *Eastern European Economics*, 58 (3), 197–220. DOI: 10.1080/00128775.2020.1726185.
- Litterman R. (1979). Techniques of forecasting using vector autoregressions. *Working Papers 115*. Federal Reserve Bank of Minneapolis.
- Litterman R. (1986). Forecasting with Bayesian vector autoregressions — five years of experience. *Journal of Business and Economic Statistics*, 4 (1), 25–38.
- Lehmann R. (2023). READ-GER: Introducing German real-time regional accounts data for revision analysis and nowcasting. *CESifo Working Paper*, 10315. DOI: 10.2139/ssrn.4394840.
- Lehmann R. (2024). A real-time regional accounts database for Germany with applications to GDP revisions and nowcasting. *Empirical Economics*, 67, 817–838. DOI: 10.1007/s00181-024-02566-3.
- Rossi B., Sekhposyan T. (2015). Macroeconomic uncertainty indices based on nowcast and forecast error distributions. *American Economic Review: Papers and Proceedings*, 105 (5), 650–655. DOI: 10.1257/aer.p20151124.
- Strohsal T., Wolf E. (2020). Data revisions to German national accounts: Are initial releases good nowcasts? *International Journal of Forecasting*, 36 (4), 1252–1259. DOI: 10.1016/j.ijforecast.2019.12.006.

Поступила в редакцию 10.12.2024;
принята в печать 10.03.2025

Приложение

Таблица П1. Зависимые переменные и календарь их выхода
(все переменные используются в темпах роста)

Переменная	Описание	Источник	Календарь выхода
<i>GDP</i>	ВВП России по использованию, % г/г	СЭПР Росстата	Первая ревизия выходит с задержкой 2 месяца, далее корректируется
<i>House_Con</i>	Расходы на конечное потребление домашних хозяйств, % г/г	СЭПР Росстата	Первая ревизия выходит с задержкой 3 месяца, далее корректируется

Окончание табл. III

Переменная	Описание	Источник	Календарь выхода
<i>Gov_Cons</i>	Расходы на конечное потребление государственного управления, % г/г		
<i>Non-com_Cons</i>	Расходы на конечное потребление некоммерческих организаций, обслуживающих домашние хозяйства, % г/г		
<i>Exports</i>	Экспорт, % г/г		
<i>Imports</i>	Импорт, % г/г		

Таблица П2. Объясняющие переменные и календарь их выхода

Переменная	Описание	Источник	Календарь выхода
<i>Unempl</i>	Уровень зарегистрированной безработицы	СЭПР Росстата	5 число месяца, задержка 1 месяц
<i>Wage</i>	Средняя номинальная начисленная заработная плата, % г/г		
<i>Credits</i>	Общий объем предоставленных средств в кредит, % г/г		
<i>USD</i>	Динамика официальных курсов иностранных валют: доллар США, % г/г		
<i>EUR</i>	Динамика официальных курсов иностранных валют: евро, % г/г		
<i>Wholesale</i>	Оборот оптовой торговли, % г/г		
<i>Servises</i>	Объем платных услуг населению, % г/г		
<i>Catering</i>	Оборот общественного питания, % г/г		
<i>Retail_Food</i>	Оборот розничной торговли: пищевые продукты, % г/г		
<i>Retail_NonFood</i>	Оборот розничной торговли: непродовольственные товары, % г/г		
<i>Freight_Transport</i>	Грузооборот транспорта, % г/г		
<i>Public_Transport</i>	Пассажиروоборот транспорта общего пользования, % г/г		
<i>Constr</i>	Объем работ, выполненные по виду деятельности «Строительство», % г/г		
<i>House</i>	Ввод в действие жилых домов, % г/г		
<i>Agricult</i>	Производство продукции сельского хозяйства, % г/г		
<i>Manufact</i>	Обрабатывающие производства, % г/г		
<i>Mining</i>	Добыча полезных ископаемых, % г/г		
<i>Prom</i>	Индекс промышленного производства, % г/г		
<i>Basic_Output</i>	Фактический объем выпуска по базовым видам экономической деятельности, % г/г		
<i>Electr</i>	Обеспечение электрической энергией, газом и паром, кондиционирование воздуха, % г/г		
<i>Urals</i>	Спотовая цена сырой нефти Urals, % г/г	Investing.com	Ежедневные данные, без задержек
<i>NatGas</i>	Цена на фьючерс на природный газ, % г/г		
<i>PMI_Russia</i>	Индекс производственной активности PMI России, % г/г	S&P Global	5–10 число каждого месяца, без задержек

Makeeva N. M. The impact of the official statistics revision on the accuracy of the Russian macroeconomic indicators nowcasting models. *Applied Econometrics*, 2025, v. 79, pp. 27–49. DOI: 10.22394/1993-7601-2025-79-27-49

Natalia Makeeva

HSE University, Moscow, Russian Federation;
nmakeeva@hse.ru

The impact of the official statistics revision on the accuracy of the Russian macroeconomic indicators nowcasting models

The paper presents the results of an accuracy analysis of nowcasting models for Russia's GDP and its components based on usage data for the period from the first quarter of 2014 to the third quarter of 2023. The novelty of the study lies in comparing the accuracy of various models — MIDAS, MFBVAR, DFM models, regularization models, and classical autoregression for order 1, evaluated on both the first and final versions of official statistics revisions, taking into account data availability within the examined quarter. The result of our study is a strategy for optimal selection of data and models for the most accurate first and final versions of nowcasting official statistics revisions for GDP and its components by use. Our paper also demonstrates that the nowcasting model for Russia's GDP, evaluated on the first revision, is more accurate regardless of data completeness within the quarter, despite multiple subsequent revisions of initial data by Rosstat. Furthermore, for GDP models, the forecasts of the first, most volatile, and imprecise revision turn out to be the most accurate.

Keywords: nowcasting; Russian GDP; time series; short-term forecasting; data revision.

JEL classification: C53; E37.

References

- Astafieva E. V., Turuntseva M. Yu. (2021). Revisions of GDP: Data and assessment of statistical properties. *HSE Economic Journal*, 25 (1), 65–101 (in Russian). DOI: 10.17323/1813-8691-2021-25-1-65-101.
- Gareev M., Polbin A. (2022). Nowcasting Russia's key macroeconomic variables using machine learning. *Voprosy Ekonomiki*, 8, 133–157 (in Russian). DOI: 10.32609/0042-8736-2022-8-133-157.
- Zhemkov M. I. (2021). Short-term assessment of Russia's GDP using the method of combining forecasts. *Series of reports on economic research, Bank of Russia, February 2021* (in Russian).
- Zubarev A., Rybak K. (2021). GDP nowcasting: Dynamic factor model vs. official forecasts. *Russian Economic Development*, 28 (12), 34–40 (in Russian).
- Makeeva N. M., Stankevich I. P. (2022). Nowcasting of the components of Russian GDP. *HSE Economic Journal*, 26 (4), 598–622 (in Russian). DOI: 10.17323/1813-8691-2022-26-4-598-622.
- Makeeva N. M., Stankevich I. P., Lyubaykin N. S. (2024). Nowcasting the Russian economy macroeconomic indicators under uncertainty: Does taking into account the news sentiment help? *Voprosy Ekonomiki*, 3, 120–142 (in Russian). DOI: 10.32609/0042-8736-2024-3-120-142.
- Stankevich I. (2020). Comparison of macroeconomic indicators nowcasting methods: Russian GDP case. *Applied Econometrics*, 59, 113–127 (in Russian). DOI: 10.22394/1993-7601-2020-59-113-127.
- Stankevich I. (2023). Application of Markov-switching MIDAS models to nowcasting of GDP and its components. *Applied Econometrics*, 70, 122–143 (in Russian). DOI: 10.22394/1993-7601-2023-70-122-143.

- Turuntseva M. Yu. (2011). Forecasting in Russia: An overview of key models. *Ekonomicheskaya Politika*, 1, 193–202 (in Russian).
- Fokin N. (2023). Nowcasting and forecasting key Russian macroeconomic variables with the MFBVAR model. *Ekonomicheskaya Politika*, 18 (3), 110–135 (in Russian). DOI: 10.18288/1994-5124-2023-3-110-135.
- Aastveit K. A., Gerdrup K. R., Jore A. S., Thorsrud L. A. (2014). Nowcasting GDP in real time: A density combination approach. *Journal of Business and Economic Statistics*, 32 (1), 48–68. DOI: 10.1080/07350015.2013.844155.
- Anesti N., Galvão A. B., Miranda-Agrippino S. (2022). Uncertain Kingdom: Nowcasting GDP and its revisions. *Journal of Applied Economics*, 37 (1), 42–62. DOI: 10.1002/jae.2845.
- Auroba S. B. (2008). Data revisions are not well-behaved. *Journal of Money Credit and Banking*, 40, 319–340. DOI: 10.1111/j.1538-4616.2008.00115.x.
- Barbaglia L., Frattarolo L., Onorante L., Pericoli F. M., Ratto M., Pezzoli L. T. (2023). Testing big data in a big crisis: Nowcasting under Covid-19. *International Journal of Forecasting*, 39 (4), 1548–1563. DOI: 10.1016/j.ijforecast.2022.10.005.
- Berger T., Morley J., Wong B. (2023). Nowcasting the output gap. *Journal of Econometrics*, 232 (1), 18–34. DOI: 10.1016/j.jeconom.2020.08.011.
- Carriero A., Clements M. P., Galvão A. B. (2015). Forecasting with Bayesian multivariate vintage-based VARs. *International Journal of Forecasting*, 35, 757–768. DOI: 10.1016/j.ijforecast.2014.05.007.
- Cepni O., Guney I., Swanson N. (2020) Forecasting and nowcasting emerging market GDP growth rates: The role of latent global economic policy uncertainty and macroeconomic data surprise factors. *Journal of Forecasting*, 39 (1), 18–36. DOI: 10.1002/for.2602.
- Chikamatsu K., Hirakata N., Kido Y., Otaka K. (2018). Nowcasting Japanese GDPs. *Bank of Japan Working Paper Series 18-E-18*. Bank of Japan.
- Clarida R., Kitchen J., Monaco R. (2002). Real-time forecasting in practice: A description of the U.S. Treasury's real-time GDP forecasting system. *MPRA Paper 21068, University Library of Munich, Germany*.
- Clements M. P., Galvão A. B. (2009). Forecasting US output growth using leading indicators: An appraisal using MIDAS models. *Journal of Applied Econometrics*, 24 (7), 1187–1206. DOI: 10.1002/jae.1075.
- Clements M. P., Galvão A. B. (2012). Improving real-time estimates of output gaps and inflation trends with multiple-vintage VAR models. *Journal of Business and Economic Statistics*, 30 (4), 554–562. DOI: 10.2307/41739785.
- Clements M. P., Galvão A. B. (2017). Predicting early data revisions to U.S. output growth and the effects of releases on equity markets. *Journal of Business and Economic Statistics*, 35 (3), 591–604. DOI: 10.1080/07350015.2015.1076726.
- Croushore D. (2008). Revisions to PCE inflation measures: Implications for monetary policy. *Working paper no. 08–8*. Federal Reserve Bank of Philadelphia.
- Dahlhaus T., Guénette J., Vasishtha G. (2017). Nowcasting BRIC+M in real time. *International Journal of Forecasting*, 33 (4), 915–935. DOI: 10.1016/j.ijforecast.2017.05.002.
- Doz C., Giannone D., Reichlin L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, 164 (1), 188–205. DOI: 10.1016/j.jeconom.2011.02.012.
- Evans M. D. D. (2005). Where are we now? *International Journal of Central Banking*, 1, 127–175.

- Galvão A. B. (2017). Data revisions and DSGE models. *Journal of Econometrics*, 196 (1), 215–232. DOI: 10.1016/j.jeconom.2016.09.006.
- Galvão A. B., Mitchell J. (2019). Measuring data uncertainty: An application using the Bank of England's "fan charts" for historical GDP growth. *EMF Working Paper Series*, 24.
- Giannone D., Reichlin L., Small D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55 (4), 665–676.
- Giovannelli A., Proietti T., Citton A., Ricchi O., Tegami C., Tinti C. (2020). Nowcasting GDP and its components in a data-rich environment: The merits of the indirect approach. *CEIS Research Paper*, 489. DOI: 10.2139/ssrn.3614110.
- Ghysels E., Kvedaras V., Zemlys V. (2016). Mixed frequency data sampling regression models: The R package midasr. *Journal of Statistical Software, Foundation for Open Access Statistics*, 72, 1–35. DOI: <http://hdl.handle.net/10.18637/jss.v072.i04>.
- Hayashi F., Tachi Y. (2021). Generalized nowcast revision analysis. DOI: 10.2139/ssrn.3904418.
- Hayashi F., Tachi Y. (2023). Nowcasting Japan's GDP. *Empirical Economics*, 64 (4), 1699–1735. DOI: 10.1007/s00181-022-02301-w.
- Howrey E. P. (1978). The use of preliminary data in economic forecasting. *The Review of Economics and Statistics*, 60 (2), 193–200.
- Howrey E. P. (1984). Data revisions, reconstruction and prediction: An application to inventory investment. *The Review of Economics and Statistics*, 66 (3), 386–393.
- Jacobs J. P. A. M., van Norden S. (2011). Modeling data revisions: Measurement error and dynamics of 'true' values. *Journal of Econometrics*, 161 (2), 101–109. DOI: 10.1016/j.jeconom.2010.04.010.
- Jardet C., Meunier B. (2022). Nowcasting world GDP growth with high-frequency data. *Journal of Forecasting*, 41 (6), 1181–1200. DOI: 10.1002/for.2858.
- Kishor N. K., Koenig E. F. (2012). VAR estimation and forecasting when data are subject to revision. *Journal of Business and Economic Statistics*, 30 (2), 181–190. DOI: 10.1198/jbes.2010.08169.
- Kočenda E., Poghosyan K. (2020). Nowcasting real GDP growth: Comparison between old and new EU countries. *Eastern European Economics*, 58 (3), 197–220. DOI: 10.1080/00128775.2020.1726185.
- Litterman R. (1979). Techniques of forecasting using vector autoregressions. *Working Papers 115*. Federal Reserve Bank of Minneapolis.
- Litterman R. (1986). Forecasting with Bayesian vector autoregressions — five years of experience. *Journal of Business and Economic Statistics*, 4 (1), 25–38.
- Lehmann R. (2023). READ-GER: Introducing German real-time regional accounts data for revision analysis and nowcasting. *CESifo Working Paper*, 10315. DOI: 10.2139/ssrn.4394840.
- Lehmann R. (2024). A real-time regional accounts database for Germany with applications to GDP revisions and nowcasting. *Empirical Economics*, 67, 817–838. DOI: 10.1007/s00181-024-02566-3.
- Rossi B., Sekhposyan T. (2015). Macroeconomic uncertainty indices based on nowcast and forecast error distributions. *American Economic Review: Papers and Proceedings*, 105 (5), 650–655. DOI: 10.1257/aer.p20151124.
- Strohsal T., Wolf E. (2020). Data revisions to German national accounts: Are initial releases good nowcasts? *International Journal of Forecasting*, 36 (4), 1252–1259. DOI: 10.1016/j.ijforecast.2019.12.006.

Received 10.12.2024; accepted 10.03.2025.