

Помогают ли высокочастотные данные в прогнозировании российской инфляции?*

Д. В. Третьяков, Н. Д. Фокин

Российская академия народного хозяйства и государственной службы при Президенте РФ,
Российская Федерация, 117517, Москва, пр. Вернадского, 82

Для цитирования: Третьяков Д. В., Фокин Н. Д. (2021) Помогают ли высокочастотные данные в прогнозировании российской инфляции? *Вестник Санкт-Петербургского университета. Экономика*. Т. 37. Вып. 2. С. 318–343. <https://doi.org/10.21638/spbu05.2021.206>

В связи с тем, что в конце 2014 г. Центральный банк совершил переход к новому для России режиму монетарной политики — режиму инфляционного таргетирования, задача прогнозирования темпов инфляции стала как никогда актуальной. В новом режиме денежно-кредитной политики Банку России важно как можно быстрее и точнее оценивать будущие темпы инфляции, чтобы максимально оперативно принимать меры по возвращению инфляции к целевому уровню. Кроме того, для проведения эффективной денежно-кредитной политики население должно обладать доверием к действиям монетарных властей и осознавать будущую динамику инфляции. Таким образом, Центральному банку для управления инфляционными ожиданиями экономических агентов необходимо активно использовать информационный канал, публикуя точные прогнозы темпа роста потребительских цен. Целью данной работы является построение модели для наукастинга, а также краткосрочного прогнозирования российской инфляции с использованием высокочастотных данных. Их включение в модель для прогнозирования весьма перспективно, так как такой подход позволяет использовать на порядок больше информации о динамике макроэкономических показателей. В работе показано, что оцененные спецификации MIDAS-моделей, в которых участвуют ряды недельной частоты номинального обменного курса рубля к доллару, ставки межбанковского кредитования MIACR, а также цены на нефть, имеют более точный прогноз месячной инфляции по сравнению с несколькими базовыми моделями, в которых используются лишь низкочастотные, то есть месячные, данные.

Ключевые слова: инфляция, наукастинг, прогнозирование, высокочастотные данные, MIDAS-модель.

Введение

Прогнозирование различных макроэкономических показателей является важнейшей задачей для центральных банков, финансовых учреждений и других институтов, для оптимального функционирования которых важно осознание будущей макроэкономической ситуации. В частности, прогнозирование инфляции является важной задачей для стран, которые стремятся ее таргетировать. В конце 2014 г. Центральный банк РФ совершил переход от режима управляемого номи-

* Статья подготовлена в рамках выполнения научно-исследовательской работы государственного задания РАНХиГС.

нального курса рубля к режиму таргетирования инфляции, целью которого является удержание темпа инфляции на постоянном, установленном заранее уровне. В нашей стране таргет, как известно, составляет 4%. Качественный прогноз темпа инфляции, сделанный до публикации официальной статистики, позволит Центральному банку проводить более правильную монетарную политику, адекватно оценивать текущее состояние экономики и своевременно принимать эффективные меры по воздействию на него. Именно поэтому задача прогнозирования инфляции настолько актуальна в России в настоящее время.

Прогнозирование инфляции является достаточно сложной задачей. Например, в статье [Stock, Watson, 2007] на американских данных приводятся свидетельства того, что наиболее качественные прогнозы инфляции получаются на основе однофакторных моделей или моделей скользящих средних. Так, в данной работе лучшей из рассмотренных оказалась модель интегрированного процесса скользящего среднего с изменяющимися во времени параметрами. Модели, использующие дополнительные факторы, продемонстрировали слабую прогнозную силу. То есть лучший прогноз был получен на использовании лишь прошлых данных о динамике самой инфляции, а точнее, ее прошлых шоков.

Согласно современной постановке неокейнсианской кривой Филлипса [Galí, Gertler, 1999], инфляция также зависит от ожидаемых значений, а не только от собственных лагов (случай адаптивных ожиданий) и разрыва выпуска или безработицы. Таким образом, текущая инфляция определяется в том числе ожиданиями экономических агентов относительно ее будущей динамики. В связи с этим, если домашние хозяйства и фирмы будут иметь завышенные инфляционные ожидания, удержание инфляции на целевом уровне Банком России будет осложнено. Это еще одна причина, по которой Центральному банку важно публиковать достоверный прогноз о будущей динамике инфляции. Конечно, не все агенты имеют рациональные ожидания, поэтому в данном случае влияние Центрального банка через информационный канал ограничено, так как повлиять на агентов, имеющих адаптивные ожидания, через данный канал он не в силах. Однако, как показано в работе [Зубарев, 2018] об оценке гибридной кривой Филлипса для России, рациональные ожидания преобладают над адаптивными. В оцененных спецификациях коэффициент при ожидаемой инфляции в целом больше, чем при лаге инфляции. Кроме того, автор отмечает, что, скорее всего, инфляционные ожидания российских агентов имеют тенденцию к рационализации, в связи с чем Банку России необходимо как можно активнее использовать информационный канал для управления инфляционными ожиданиями.

Насколько нам известно, в академической литературе представлено не так много работ, посвященных прогнозированию российской инфляции. В работе [Андреев, 2016] описывается один из методов, которым пользуется Центральный банк при краткосрочном прогнозировании, — метод комбинирования прогнозов. Данный метод основан на прогнозировании субиндексов индекса потребительских цен (ИПЦ) и использует несколько эконометрических моделей, таких как случайное блуждание, авторегрессия с детерминированным трендом, модель ненаблюдаемых компонент, ARX-модель, VAR- и BVAR-модели по разным комбинациям переменных. Затем строится линейная комбинация прогнозов с весами, пропорциональными их качеству. После чего прогнозы по субиндексам ИПЦ аналогично

взвешиваются, и получается итоговый прогноз ИПЦ. Помимо этого, для решения задачи прогнозирования инфляции в литературе использовались такие модели машинного обучения, как Lasso-регрессия, случайный лес, бустинг и другие [Байбуза, 2018], нейронные сети [Павлов, 2020], а также байесовское усреднение моделей [Стырин, 2019]. В работе [Байбуза, 2018] была показана состоятельность моделей машинного обучения, то есть их способность строить более качественные, чем традиционные, эконометрические модели, прогнозы. В исследовании [Павлов, 2020] было продемонстрировано превосходство простой нейронной сети в прогнозировании месячной инфляции относительно стандартного бенчмарка AR(1). В работе [Стырин, 2019] было показано, что байесовское усреднение проигрывает более простым моделям. Метод, используемый в последней работе, в каком-то смысле отбирает наиболее релевантные регрессоры для прогноза из достаточно большого набора данных. Таким образом, был получен результат, в чем-то аналогичный результату из упомянутой работы [Stock, Watson, 2007], но на российских данных.

Для прогнозирования инфляции в зарубежной литературе применяются различные модели эконометрики и машинного обучения. Отметим несколько недавно опубликованных работ. В работе [Delle Monache, Petrella, 2017] предлагается алгоритм построения векторной авторегрессии с меняющимися во времени параметрами при наличии тяжелых хвостов. Авторы заключают, что предложенная модель дает более качественные прогнозы инфляции в США и странах Большой семерки по сравнению с бенчмарками. В работе [Medeiros et al., 2021] демонстрируется высокое качество модели случайного леса в точности прогнозов американской инфляции. Данная модель способна выбрать полезные предикторы из их большого числа и учесть нелинейные взаимосвязи между инфляцией и ее детерминантами. Также в литературе представлены структурные эконометрические модели, которые выводятся из экономической теории. Например, в работе [McKnight, Mihailov, Rumler, 2020] авторы прогнозируют инфляцию США и еврозоны на основе неокейнсианской кривой Филлипса. В модели предполагается меняющийся во времени трендовый уровень инфляции с целью моделирования структурных сдвигов в денежно-кредитной политике центральных банков.

При этом ни в одной из упомянутых отечественных работ не ставилась задача наукастинга (nowcasting¹) инфляции. Под термином «наукастинг» понимается фактически оценка значения объясняемой переменной в текущем периоде, когда некоторые данные о регрессорах в правой части уже известны, а значение объясняемой переменной — еще нет². Фактически актуальность наукастинга рождается из-за технических особенностей процесса сбора и публикации данных. Например, биржевые цены доступны каждую минуту, и даже чаще. Они очевидным образом могут влиять на целый ряд макроэкономических показателей, статистика по которым достаточно сложно собирается и которые долго рассчитываются.

¹ Одной из первых экономических работ, где употреблялся этот термин, является исследование [Baffigi, Golinelli, Parigi, 2004].

² В данной работе используется следующая терминология по аналогии со стандартными терминами в прогнозировании временных рядов. Термин «наукастинг» (nowcasting) — по смыслу аналог слова «прогнозирование» (forecasting), под которым понимается процесс построения прогнозных значений временного ряда. Термин «наукаст» (nowcast) — по смыслу аналог слова «прогноз» (forecast), то есть результат наукастинга/прогнозирования.

Официальная статистика по инфляции за месяц t выходит в начале следующего месяца $t + 1$. Предположив, например, связь между обменным курсом рубля к доллару, доступным каждый день, и инфляцией, мы можем оценить значение инфляции (построить наукаст) в месяце t на основе дневных данных по номинальному обменному курсу в этом месяце. Такая оценка, начиная с первого числа месяца t , и называется наукастом. С увеличением числа дневных наблюдений номинального курса наукаст может становиться точнее. Кроме того, по сравнению с использованием классических регрессионных моделей данных одинаковой частоты использование дополнительных данных о динамике высокочастотных (дневных) регрессоров внутри низкочастотного периода (месяца) может увеличить прогнозную силу модели, за счет того, что увеличивается число наблюдений.

В данной работе используются две модели, связывающие высокочастотные и низкочастотные данные. Они будут подробно описаны ниже, однако краткое описание дадим сейчас. Первая, так называемая модель мостовых уравнений (Bridge equations, Bridge), предлагалась фактически впервые в работах [Trehan, 1989; Parigi, Schlitzer, 1995]. Ее суть состоит в приведении высокочастотных данных в правой части к аналогичной частоте с низкочастотными в левой части путем простого усреднения. По аналогии с примером инфляции и номинального курса в самом простом случае оценивается регрессия инфляции на номинальный курс на месячных данных. Для построения наукаста на месяц t номинальный курс усредняется по дневным наблюдениям за месяц t , и полученное значение используется в качестве значения месячного номинального курса. По полученным оценкам регрессии рассчитывается наукаст.

Вторая модель — MIDAS (MIXed DATA Sampling), предложенная в работе [Ghysels, Santa-Clara, Valkanov, 2004], — основана уже не на простом усреднении высокочастотных данных в низкочастотные, а на непосредственном использовании высокочастотных данных и их динамики внутри низкочастотного периода. Самая базовая ее версия, именуемая U-MIDAS, является линейной регрессией низкочастотной переменной на несколько высокочастотных переменных, количество которых равно числу периодов высокой частоты в низкочастотной переменной. В случае с месячными и дневными данными, если предположить, что в каждом наблюдаемом месяце 30 дней, мы будем регрессировать низкочастотную месячную переменную на 30 дневных переменных. В первой переменной будут содержаться первые дни всех месяцев, во второй — вторые дни всех месяцев и т. д. В данном случае очевидно усугубление проклятия размерности с ростом разницы в частоте данных, в связи с чем существует несколько иных спецификаций MIDAS-моделей, предназначенных для избежания излишней параметризации. В данной работе мы в основном концентрируемся на построении MIDAS-моделей, в связи с тем, что они новее и являются более гибкими, чем Bridge. Забегая вперед, отметим, что MIDAS-модели продемонстрировали более высокую прогнозную силу. Кроме наукастинга, в работе также тестируется качество краткосрочных прогнозов (на 1–2 месяца) на основе MIDAS-, Bridge-моделей, а также рассмотренных бенчмарков: среднего, наивного прогноза и ARIMA-модели. Отметим, что в российской литературе подход MIDAS использовался, например, для прогнозирования ВВП [Микош, Соланко, 2019]. В данной работе авторы получили более высокое качество прогнозов по MIDAS-модели, чем по моделям Bridge и авторегрессионной модели на периоде

2008–2011 гг. Существуют также эмпирические работы в зарубежной литературе, в которых решалась задача, аналогичная нашей, и использование MIDAS-модели существенно улучшало качество наукаста и прогнозов инфляции [Ribon, Suhoj, 2011; Breitung, Røling, 2015].

Таким образом, новизной данного исследования является решение новой задачи в области прогнозирования темпа инфляции, а именно построение модели, позволяющей оценивать текущее значение инфляции (наукаст), а также использование альтернативного способа краткосрочного прогнозирования инфляции — с помощью высокочастотных данных.

Работа построена следующим образом. В первом разделе рассматриваются основные методы и модели, которые использовались для прогнозирования инфляции, а также выделяются ее основные детерминанты, применяемые в теоретических и эмпирических моделях. Во втором разделе описывается методология MIDAS- и Bridge-моделей. В третьем разделе описываются используемые данные и формулируются эксперименты по тестированию прогнозной силы моделей. В четвертом, последнем разделе приведены основные результаты работы.

1. Моделирование инфляции. Обзор подходов и методов

Детерминанты инфляции со временем могут изменяться. Так, может меняться характер инфляции с немонетарного на монетарный. Может меняться структура инфляционных ожиданий, а также режим денежно-кредитной политики (ДКП), как это случилось в России в конце 2014 г. Следовательно, необходимо внимательно подойти к отбору факторов, которые потенциально могут быть полезными для задачи прогнозирования инфляции. В этом разделе представлены работы, направленные на изучение факторов, объясняющих инфляцию в России и некоторых других странах, а также предлагающие различные подходы к моделированию инфляции как в условиях инфляционного таргетирования, так и в режиме управляемого номинального обменного курса.

Для проведения Центральным банком эффективной ДКП важно понимать, каким образом те или иные факторы оказывают влияние на уровень инфляции в стране. Так, в случае режима плавающего валютного курса колебания обменного курса национальной валюты более сильны, чем при режиме управляемого курса, и поэтому инфляция, которая считается на основе ИПЦ, в большей степени подвержена изменениям за счет изменения цен на импортные товары. Эффект влияния обменного курса на уровень цен именуется эффектом переноса обменного курса в цены. В России данный эффект был достаточно подробно изучен в статье [Пономарев, Трунин, Улюкаев, 2014]. В данной работе было получено наличие значимого эффекта переноса курса в цены в России. Авторы оценивали VECM-модель и получили наличие статистически значимого неполного эффекта переноса (эластичность уровня цен по обменному курсу меньше единицы) для периода 2000–2012 гг. Авторы выделяют три основных механизма изменения инфляции за счет обменного курса: 1) прямой эффект — за счет удорожания импортных товаров; 2) косвенный эффект, который состоит в росте спроса со стороны иностранных потребителей за счет удешевления отечественных товаров; и 3) эффект прямых иностранных инвестиций. Если национальная валюта страны испытывает продол-

жительное ослабление, то зарубежные производители в целях минимизации издержек могут перенаправлять производство в эту страну.

В работе [Пономарев, Трунин, Улюкаев, 2014] было показано, что эффект переноса обменного курса евро на потребительские цены чаще ниже, чем доллара США. Была получена оценка эффекта переноса в России, которая больше схожа с показателями развивающихся стран, для которых характерен более высокий уровень инфляции по сравнению с развитыми странами. Учитывая переход к режиму свободного курса, очевидно, что номинальный курс и сейчас остается одной из важнейших детерминант российской инфляции.

Одним из основных инструментов, с помощью которого Центральный банк воздействует на инфляцию, является ключевая ставка. Процентный канал трансмиссионного механизма ДКП работает в два основных этапа. Первый связан непосредственно с действиями Центрального банка, когда он с помощью доступных инструментов воздействует на ставку процента на рынке межбанковского кредитования. Эти действия вызывают изменения в краткосрочных процентных ставках денежного рынка, то есть происходит изменение в доходности государственных облигаций, цен других активов. Вследствие этого происходит перенос изменений в краткосрочные и долгосрочные ставки финансового рынка. Второй этап связан с влиянием изменения ставок на поведение экономических агентов, то есть на их решения об объеме заимствований и сбережений, производства и потребления, то есть на реальный сектор экономики. Таким образом, происходит корректировка спроса в экономике и уровня цен.

Если изменение процентной ставки межбанковского кредитования быстро и полностью переносится на ставки по кредитам банков, то монетарная политика будет эффективной, а значит, произойдет быстрое влияние на внутренний спрос и инфляцию. Основными причинами неполноты, слабости эффекта переноса ставок межбанковского рынка на ставки банковского сектора являются асимметрия информации на рынке, концентрация банковской системы, издержки меню, издержки переключения (связаны с переходом от одного банка к другому). Подробнее с данным эффектом можно ознакомиться в работе [Перевышин, Перевышина, 2015].

Как было замечено во введении, инфляционные ожидания являются детерминантой текущей инфляции, если следовать современным версиям кривой Филлипса. Согласно теории рациональных ожиданий [Lucas, 1973; Sargent, Wallace, 1975; 1976], экономические агенты при принятии решений ориентируются на всю доступную им информацию, в том числе на ожидаемую будущую динамику макроэкономических показателей. В терминах эконометрики математическое ожидание от будущего значения объясняемой переменной участвует в правой части регрессии. Использование инфляционных ожиданий в эконометрическом анализе сопряжено с несколькими проблемами. Основная проблема состоит в том, что они достаточно сложно собираются, с помощью опросов, которые проводятся по заказу Банка России. Первый такой опрос датируется мартом — апрелем 2013 г. Статистические данные по инфляционным ожиданиям публикуются с января 2016 г. каждый месяц, то есть временные ряды являются не слишком протяженными во времени, а также низкочастотными. Очевидно, слишком затратным является сбор высокочастотных данных по инфляционным ожиданиям с помощью опросов.

Отметим, что изменение цен на нефть или других экспортируемых из России товаров может приводить к росту внутренних цен по нескольким причинам. Так, повышение цен на нефть может трактоваться как трансфер богатства в отечественную экономику из-за рубежа, что приводит к повышению спроса за счет роста дохода и затем к повышению цен. Кроме того, нефть является основополагающим сырьем при производстве большого количества отечественных товаров, начиная с разных видов топлива и заканчивая пластмассой и резиной. Цены на нефть использовались в качестве детерминанты инфляции при оценке кривой Филлипса с адаптивными ожиданиями в работе [Гафаров, 2011]. Кроме того, цены на нефть являются биржевым товаром, и мы можем наблюдать данный показатель в реальном времени в очень высокой частоте.

Перейдем к обзору работ, посвященных прогнозированию инфляции. В статье [Байбуза, 2018] рассмотрены различные методы машинного обучения для решения данной задачи. Рассматривались следующие модели: регрессии Lasso, Ridge, Elastic Net, случайный лес и бустинг. Для многих моделей машинного обучения, а тем более для моделей с регуляризацией, таких как Lasso, Ridge и Elastic Net, выбор стартового набора регрессоров не является критически важным, так как в такие модели можно включить сколь угодно большое число объясняющих переменных и выбрать из них наиболее релевантные, оценив коэффициенты при них. Иными словами, данные модели не подвержены проклятию размерности. В качестве регрессоров выбирались такие, которые отражают состояние денежного рынка, промышленного производства, уровня занятости и платежного баланса, а также цен на товары, которые экспортируются Россией. Ряд ИПЦ берется за период с февраля 2002 г. по январь 2018 г. В качестве основных бенчмарков автор выбрал AR(1)- и AR(p)-модели, где глубина запаздываний выбирается на основе критерия Шварца. Также рассматривалась комбинация AR(1) и Lasso (то есть сперва регрессоры отбираются с помощью Lasso и далее они включаются вместе с первым лагом в расширенную модель, полученная спецификация оценивается методом наименьших квадратов (ordinary least squares, OLS)). Данные стандартизируются, из рядов вычитается среднее, и они делятся на стандартное отклонение. Наибольшие по модулю оценки у комбинированной модели имеют следующие регрессоры: предыдущее значение инфляции, кредиты нефинансовым организациям, объем работ по виду деятельности «строительство», ключевая ставка Банка России, оборот розничной торговли.

Статья [Стырин, 2019] посвящена использованию метода динамического усреднения моделей (Dynamic Model Averaging) для прогнозирования инфляции. В данной работе инфляция рассматривается с января 2002 г. по сентябрь 2017 г. В качестве бенчмарков использовались: байесовское усреднение моделей (Bayesian Model Averaging), байесовский выбор модели (Bayesian Model Selection), модель с ненаблюдаемыми компонентами и стохастической волатильностью, байесовская авторегрессия второго порядка с постоянными и изменяющимися во времени коэффициентами. На основе результатов псевдовневыборочного прогнозирования на 1–6 месяцев рассматриваемая автором в качестве основной динамически усредненная модель продемонстрировала более слабую прогнозную силу, чем рассматриваемые бенчмарки. Самую высокую апостериорную вероятность включения в прогнозную модель имеют кредиты физическим лицам и организациям, а также ожидаемые и фактические изменения заработной платы.

В статье [Павлов, 2020] автор сравнивает прогнозную силу простейшей нейронной сети с одним скрытым слоем, метода опорных векторов и Ridge-регрессии с эталонной однофакторной моделью AR(1) применительно к месячной российской инфляции. В качестве предикторов использовались 10 основных макроэкономических показателей: реальный валовой внутренний продукт (ВВП), агрегат денежной массы M2, объем выданных кредитов в реальном выражении, процентная ставка денежного рынка и др. Согласно результатам, нелинейные модели доминируют над линейными при увеличении горизонта прогнозирования. Так, в ситуации, когда он составляет 2, 6 и 24 месяца, нейронная сеть имеет наименьшую ошибку прогноза среди всех моделей.

В зарубежной литературе есть ряд работ на тему наукастинга инфляции. В статье [Breitung, Røling, 2015] используется регрессия MIDAS для прогнозирования инфляции в Германии на данных за период 2000–2014 гг. Обучение происходит до 2006 г., и далее строится рекурсивный прогноз с движущимся окном в один шаг. В качестве бенчмарка авторы рассматривают модель с усреднением дневных показателей в месячные (Bridge). Авторы используют в качестве регрессоров переменные дневной частоты, такие как доходности облигаций, цены на различные товары и индексы цен на акции. Конечная модель основывается на ежедневном индексе цен на сырьевые товары с использованием параметрических и непараметрических структур MIDAS. Основные выводы состоят в том, что индекс цен на товары HWWI (местный индекс, который публикуется немецким университетом Hamburgisches WeltWirtschaftsinstitut (HWWI) и отражает цены на промышленные, сырьевые товары, драгоценные металлы, а также на зерновые культуры, масла и т. д.) является полезным при прогнозировании на 1 месяц вперед. Оцененное распределение лагов имеет форму горба с пиком между 5–10 днями. Также вывод авторов состоит в том, что для прогнозирования инфляции дневными данными должны быть использованы более гибкие распределения запаздывания, то есть модель должна включать в себя полиномы с большим количеством параметров.

В работе [Ribon, Suhoу, 2011] авторы применяли модель MIDAS для прогнозирования и наукастинга инфляции в Израиле. В качестве высокочастотных (дневных) данных были выбраны номинальный обменный курс, индекс товарных цен, ключевая ставка. Также в модель были включены предыдущие значения инфляции и инфляционные ожидания. Результаты показывают, что модель MIDAS имеет бóльшую прогнозную силу по сравнению с более простыми бенчмарками. По результатам исследования MIDAS модель оказалась лучшей в прогностической силе на выборке 2008–2010 гг.

Подводя итоги обзора и забегая несколько вперед, опишем используемые в данной работе регрессоры для инфляции. Из высокочастотных переменных мы используем номинальный обменный курс и процентную ставку MIACR, так как использование ключевой ставки в эконометрических моделях затруднено тем, что у данной переменной практически отсутствует вариация, а также номинальные цены на нефть марки BRENT. Также в модели будут участвовать лаги инфляции, что может быть интерпретировано как включение в модель адаптивных инфляционных ожиданий.

2. Методология MIDAS

В данном разделе будет описана методология моделей MIDAS, которые использовались в исследовании. Впервые модель MIDAS была представлена в статье [Ghysels, Santa-Clara, Valkanov, 2004]. Позднее первоначальные идеи были развиты в работе [Ghysels, Kvedaras, Zemlys, 2016], где помимо добавления новых спецификаций модели был предложен пакет `midasr` для языка R. В этом пакете произведены оценки моделей из данной работы.

В базовом случае задачей MIDAS-регрессии является спецификация условного математического ожидания низкочастотной переменной от высокочастотных. Модель допускает в том числе включение лагов объясняемой переменной в правую часть, а также других низкочастотных переменных.

Рассмотрим регрессионную модель, в которой присутствуют низкочастотная переменная y_t и высокочастотная x_{tj} . Пусть $t = 1, \dots, T$ обозначает низкочастотный временной индекс (месячный, например), а j обозначает высокочастотный индекс (недельный, например), так как в месяце может быть от 4 до 5 недель; для простоты будем считать, что в каждом месяце 4 недели ($p = 4$). В базовой спецификации модели U-MIDAS низко- и высокочастотные переменные связаны линейной регрессией вида

$$y_t = \gamma_0 + \sum_{j=1}^p \beta_j x_{tj} + \varepsilon_t, \quad (1)$$

где ε_t — белый шум. Матрица регрессоров выглядит следующим образом:

$$X = \begin{pmatrix} x_{1,1} & \cdots & x_{1,4} \\ \vdots & \ddots & \vdots \\ x_{T,1} & \cdots & x_{T,4} \end{pmatrix}, \quad (2)$$

где $x_{t,j}$ соответствует j -й неделе в t -м месяце.

В данной модели предполагается, что месячное значение объясняемой переменной зависит линейно от каждого недельного значения регрессора. С целью избежать излишней параметризации, которая усугубляется с ростом отношения числа высокочастотных наблюдений к низкочастотным, предлагается использовать различные полиномы от некоторого числа параметров, которые задают все коэффициенты при высокочастотных переменных:

$$y_t = \gamma_0 + \gamma_1 \sum_{j=1}^p w_j(\theta_1, \dots, \theta_n) x_{tj} + \varepsilon_t. \quad (3)$$

В уравнении (3) коэффициенты перед каждым недельным регрессором определяются исходя из значения функции $w_j(\theta_1, \dots, \theta_n)$, а также коэффициента γ_1 , отражающего некоторый общий эффект изменения низкочастотной объясняемой переменной за счет изменения высокочастотной. Функция $w_j(\theta_1, \dots, \theta_n)$ может представлять собой, например, экспоненциальный полином:

$$w_j(\theta_1, \dots, \theta_n) = \frac{\exp(\theta_1 j + \dots + \theta_n j^n)}{\sum_{i=1}^p \exp(\theta_1 i + \dots + \theta_n i^n)}, \quad (4)$$

где $w_j(\theta_1, \dots, \theta_n) \in [0, 1]$. Использование такой функциональной формы позволяет сократить число оцениваемых параметров. Так, например, в работе [Ghysels, Sinko, Valkanov, 2007] показано, что использование полиномов w_j с двумя параметрами уже достаточно для качественного прогноза.

Обозначим $\gamma_1 w_j(\theta_1, \dots, \theta_n) = \beta_j$ и перепишем уравнение (3) в матричной форме. Предположим, что данные в текущем, а также предыдущем месяце имеют предсказательную силу. Это означает, что для периода t мы моделируем y_t как линейную комбинацию переменных $x_{t,1}, x_{t,2}, x_{t,3}, x_{t,4}$, наблюдаемых в месяц t , и переменных $y_{t-1}, x_{t-1,1}, x_{t-1,2}, x_{t-1,3}, x_{t-1,4}$, наблюдаемых в месяц $t-1$:

$$\begin{pmatrix} y_2 \\ \vdots \\ y_T \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \alpha_0 + \begin{pmatrix} y_1 \\ \vdots \\ y_{T-1} \end{pmatrix} \alpha_1 + \begin{pmatrix} x_{2,4} & x_{2,3} & \dots & x_{1,1} \\ x_{3,4} & x_{3,3} & \dots & x_{2,1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{T,4} & x_{T,3} & \dots & x_{T-1,1} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_8 \end{pmatrix} + \begin{pmatrix} \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{pmatrix}. \quad (5)$$

Параметры β_1, \dots, β_8 определяются параметрами полинома $\theta_1, \dots, \theta_n$. Следуя, например, рекомендации [Ghysels, Sinko, Valkanov, 2007] и используя всего два параметра полинома θ_1, θ_2 , сократим пространство параметров в 4 раза. Конечно, когда число наблюдений очень велико, разумной является оценка модели, в которой у каждого высокочастотного регрессора присутствует свой собственный параметр, а не использование ограниченных полиномиальных функций.

Отметим, что в рамках модели требуется, чтобы число высокочастотных наблюдений было кратно T (числу низкочастотных наблюдений). Для этого выше была сделана предпосылка, что в каждом месяце 4 недели. Данный факт, конечно, является недостатком MIDAS-моделей, однако единственный на данный момент способ его избежать — фиктивно приводить к кратности число высокочастотных точек в низкочастотном периоде. В целом данный подход является стандартным в литературе (см., например: [Baumeister, Guérin, Kilian, 2015]).

Вернемся к сделанному ранее обозначению коэффициентов перед лагами высокочастотных $\beta_j = \gamma_1 w_j(\theta_1, \dots, \theta_n)$. В некоторых полиномах, например Алмона, коэффициента общего влияния γ_1 не предусмотрено. Значение второго множителя для любого из рассматриваемых полиномов можно представить в виде

$$w_j(\theta_1, \dots, \theta_n) = \frac{\Psi(\theta, j)}{\sum_{i=1}^p \Psi(\theta, i)}. \quad (6)$$

Рассмотрим различные спецификации функции $\Psi(\theta, j)$.

В уравнении (7) представлен экспоненциальный полином Алмона (Nealmon) от n параметров $\theta_1, \dots, \theta_n$

$$\Psi(\theta_1, \dots, \theta_n, j) = \exp\left(\sum_{i=1}^n \theta_i j^i\right). \quad (7)$$

В уравнении (8) представлен бета-полином (Nbeta) от двух параметров θ_1, θ_2 .

$$\Psi(\theta_1, \theta_2, j) = x_j^{\theta_1 - 1} (1 - x_j)^{\theta_2 - 1}, \quad (8)$$

где $x_j = \varepsilon + (1 - \varepsilon)h(j)$, $h(j) = \frac{j-1}{p-1}$, $\varepsilon > 0$ необходимо, чтобы x_j отличался от нуля.

В уравнении (9) представлен Log-Cauchy полином от двух параметров θ_1, θ_2 .

$$\Psi(\theta_1, \theta_2, j) = j^{-1} \left(\theta_2^2 + (\log(j) - \theta_1)^2 \right)^{-1}. \quad (9)$$

Для полинома Алмона (Almon) формула выглядит следующим образом:

$$w_j(\theta_1, \dots, \theta_n) = \sum_{i=1}^n \theta_i j^{i-1}. \quad (10)$$

Таким образом, оптимизационная задача в рамках MIDAS-регрессии сводится к оценке единого для всех высокочастотных лагов параметра θ (или двух параметров θ_1, θ_2). Изменение коэффициентов при увеличении лага происходит за счет изменения номера лага j . Модель может быть оценена нелинейным методом наименьших квадратов, путем поиска такого значения θ (и прочих параметров, если они присутствуют, например константы, параметров при экзогенных переменных), которое минимизирует остаточную сумму квадратов.

Различные полиномы способны задавать разное поведение коэффициентов перед высокочастотными регрессорами с увеличением лага. Например, в случае экспоненциального полинома Алмона коэффициенты могут плавно или быстро уменьшаться (увеличиваться) с ростом лага, а также изменяться по форме параболы. При этом они не могут быть отрицательными. То есть использование данного полинома представляется разумным в рамках предпосылки о том, что изменение недельного регрессора положительно влияет на целевую переменную.

Также в исследовании рассматривается модель Bridge [Trehan, 1989; Parigi, Schlitzer, 1995; Baffigi, Golinelli, Parigi, 2004; Angelini, Bańbura, Rünstler, 2010; Breitung, Røling, 2015], которая является регрессией низкочастотной переменной на высокочастотные переменные, усредненные в размерность низкочастотной. При наукастинге с помощью данной модели известные на момент принятия прогноза высокочастотные наблюдения усредняются и полученное значение используется для построения наукаста по оцененной линейной модели на низкочастотных данных.

3. Построение MIDAS-модели для прогнозирования российской инфляции

3.1. Описание данных и постановка модели

В качестве прогнозируемой переменной, отражающей рост общего уровня цен, мы рассматриваем первую разность логарифма месячного индекса потребительских цен. Ряд ИПЦ взят с сайта Росстата с января 2010 г. до ноября 2019 г.³ Ле-

³ Федеральная служба государственной статистики. URL: <https://rosstat.gov.ru/price> (дата обращения: 08.06.2021).

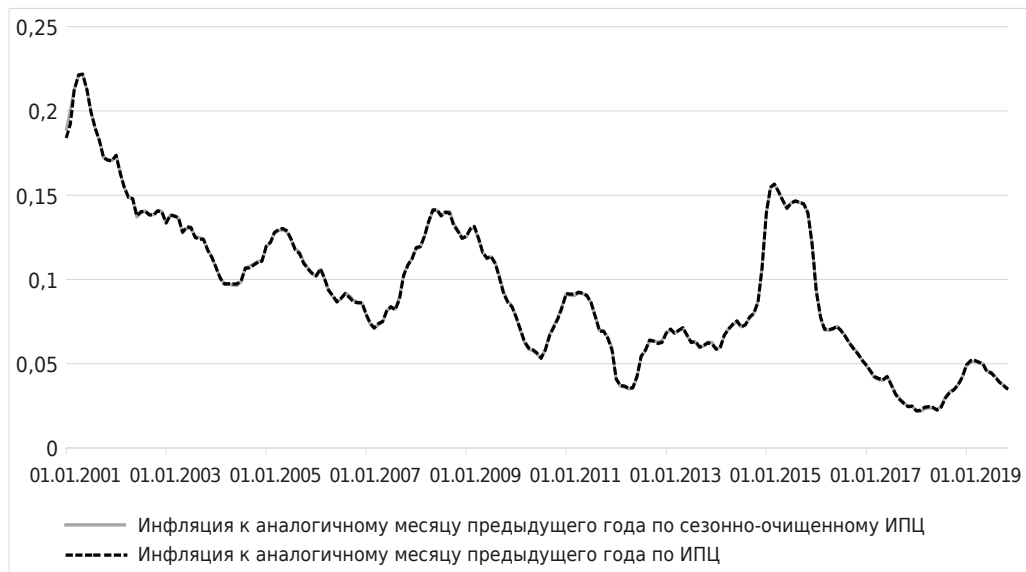


Рис. 1. Инфляция к аналогичному месяцу предыдущего года по «сырому» и сезонно-очищенному ИПЦ

вая граница рассматриваемой выборки обусловлена несколькими причинами. Из-за смены режима ДКП на режим инфляционного таргетирования в конце 2014 г., образование цен могло существенно измениться. В связи с этим при достаточном количестве точек корректно рассматривать выборку начиная с ноября 2014 г. Однако в оцениваемых в данной работе спецификациях MIDAS-моделей присутствует большое число параметров, в связи с чем мы решили несколько увеличить число наблюдений и рассматривать данные по инфляции с 2010 г. Кроме того, в то время Банк России уже отказался от жестко фиксированного обменного курса и ввел режим валютного коридора. Несколько позже появилась и ключевая ставка, и процентные ставки стали использоваться как основной инструмент ДКП.

Для получения ряда инфляции мы берем первую разность логарифма ИПЦ, очищенного от сезонности процедурой ARIMA X-12. Очистка производится на полной выборке с 2010 по 2019 г. Такой подход может быть подвержен определенной критике, так как при очистке от сезонности используются в том числе наблюдения тестового периода. Однако рассматриваемые бенчмарки мы строим на таком же временном ряде инфляции, поэтому все модели в прогнозном эксперименте находятся в одинаковых условиях. Задачу прогнозирования инфляции с учетом сезонности мы оставляем для будущих исследований, однако заметим, что полученные прогнозы инфляции к прошлому месяцу мы можем перевести в ряд инфляции к аналогичному месяцу предыдущего года. В годовом выражении ряд инфляции получается практически идентичным как для очищенного от сезонности, так и для «сырого» ИПЦ (рис. 1).

В качестве регрессоров мы используем процентную ставку MIACR, номинальный обменный курс рубля к доллару и номинальную цену на нефть. Дневные ряды номинального обменного курса рубля к доллару и ставки MIACR сроком на один

день были взяты с сайта Банка России⁴. Дневной ряд цены на нефть был взят с сайта FRED с тикером DCOILBRENTU⁵. В данной работе мы рассмотрим недельные данные в качестве высокочастотных. Мы не рассматриваем дневные данные по причине того, что их достаточно сложно обрабатывать. Так или иначе придется отбрасывать лишние наблюдения за счет разного количества бизнес-дней в месяцах и различий в выходных и праздничных днях в России и мире. Из исходных дневных данных недельные получаются путем усреднения. Мы предполагаем, что в каждом месяце есть 4 условных недели. В данном случае мы следовали подходу, использованному в статье [Baumeister, Guérin, Kilian, 2015]. Первая условная неделя заканчивается 5-м бизнес-днем, вторая условная неделя — 10-м, третья неделя примерно 15-м бизнес-днем месяца, а четвертая условная — последним днем месяца.

Перед записью модели протестируем порядок интегрированности рассматриваемых переменных. Мы будем использовать два достаточно популярных теста, а именно ADF- и KPSS-тесты. Использование такой пары тестов с разными нулевыми гипотезами в каком-то смысле позволяет нам избежать проблемы низкой мощности, за счет которой какой-то тест может не отвергать нулевую гипотезу чаще, чем положено. В ADF-тесте лаги выбираются на основе критерия Шварца. Для ряда логарифма ИПЦ в уравнении участвует константа. Недельные ряды логарифма обменного курса, логарифма цены на нефть и MIACR тестируются без константы. В KPSS-тесте используется ядро Бартлетта, а ширина окна выбирается по методу Ньюи — Веста. Результаты представлены в табл. 1.

Таблица 1. Результаты тестирования порядков интегрированности

Переменная	Статистика ADF-теста в уровнях	Статистика ADF-теста в первых разностях	Статистика KPSS-теста в уровнях	Статистика KPSS-теста в первых разностях
Логарифм ИПЦ	-1,41	-4,58***	1,29***	0,33
Логарифм обменного курса	1,18	-11,82***	2,39***	0,09
MIACR	-0,39	-21,54***	1,30***	0,14
Логарифм цен на нефть	-0,29	-17,39***	1,40***	0,10

Примечание. Здесь и в табл. 2, 6: *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$.

Как видно из таблицы, нулевая гипотеза о наличии единичного корня не отвергается тестом ADF для всех рассматриваемых рядов. При этом она отвергается в их первых разностях. Аналогичный результат получен и с помощью KPSS-теста. Данный тест отвергает гипотезу о стационарности для всех рассматриваемых рядов, но не отвергает гипотезу о стационарности для их первых разностей.

Согласно результатам проведенных тестов, мы будем рассматривать ряды логарифма ИПЦ, ставки MIACR, логарифма цен на нефть и логарифма долларо-

⁴ Банк России. URL: https://www.cbr.ru/hd_base/ (дата обращения: 08.06.2021).

⁵ FRED Economic Data. URL: <https://fred.stlouisfed.org/series/DCOILBRENTU> (дата обращения: 08.06.2021).

го курса как стационарные в их первых разностях. Наконец, оцениваемую модель MIDAS можно записать в следующем виде:

$$\pi_t = \alpha_0 + \sum_{i=1}^{p^n} \alpha_i \pi_{t-i} + \beta^e \sum_{i=0}^{p^e} \sum_{j=1}^{q_i^e} w^e(\theta^e) \Delta \ln e_{t-i,j} + \beta^m \sum_{i=0}^{p^m} \sum_{j=1}^{q_i^m} w^m(\theta^m) \Delta m_{t-i,j} + \beta^b \sum_{i=0}^{p^b} \sum_{j=1}^{q_i^b} w^b(\theta^b) \Delta \ln b_{t-i,j} + \varepsilon_t, \quad (11)$$

где π_t — первая разность логарифма, очищенного от сезонности базисного ИПЦ; $\Delta \ln e_t$ — первая разность логарифма номинального обменного курса; Δm_t — ставка MIACR в первых разностях; $\Delta \ln b_t$ — первая разность логарифма цены на нефть. Векторы θ^e , θ^m , θ^b являются векторами параметров полиномов для обменного курса, процентной ставки и цены на нефть соответственно. p с индексом обозначает число месячных лагов соответствующей переменной, q_i обозначает число недельных лагов определенной переменной в месяце $t - i$.

3.2. Процедура выбора лучшей модели для прогнозирования

В данной работе для тестирования прогнозной силы моделей будет произведен ряд экспериментов по псевдовневыборочному прогнозированию инфляции. Первым шагом построения прогноза является наукастинг. Как уже отмечалось ранее, публикация официальной инфляции за текущий месяц происходит в течение следующего. Будем предполагать, что месячные данные по инфляции нам известны с начала второй недели месяца, следующего за отчетным⁶. Данное предположение нам нужно, чтобы определять, на какое количество недель мы опережаем выход официальных данных. Рассмотрим следующие ситуации: известны все данные за текущий месяц (наукаст опережает выход официальной статистики на 1 неделю), известны только 3 недели текущего месяца (наукаст опережает выход статистики на 2 недели), известны 2 недели текущего месяца (наукаст опережает выход статистики на 3 недели) и известна 1 неделя текущего месяца (наукаст опережает выход статистики на 4 недели).

В модели (11) нужно оценить большое число параметров, выбрать глубину месячных запаздываний инфляции (от 0 до 12), а также выбрать лучшие по степени прогнозной силы полиномы (Almon, Nealmon, Log-Cauchy, Nbeta) для всех регрессоров. Для этого мы используем следующую процедуру оценки. Мы делим выборку на тренировочную и тестовую. В тренировочную выборку входят наблюдения за период 2010M1–2017M12, в тестовую — 2018M1–2019M11⁷. На тренировочной выборке оцениваются параметры уравнения и выбирается глубина запаздываний регрессоров правой части на основе информационного критерия Шварца. Так делается для всех возможных комбинаций полиномиальных функций у темпов ро-

⁶ Официально данные публикуются на шестой рабочий день. См.: Индексы потребительских цен и средние цены на товары и услуги. URL: <https://rosstat.gov.ru/storage/mediabank/ipc.html> (дата обращения: 08.06.2021).

⁷ Мы также рассмотрели модели, обученные на большей и меньшей выборках, а именно на выборках с 2001M01 по 2017M12 и с 2014M01 по 2017M12, с целью анализа стабильности результатов. Мы получили, что при наукастинге лучшей оказывается основная рассматриваемая модель, оцениваемая с 2010 г. При прогнозировании на один и два месяца вперед в ряде случаев прогноз оказывается несколько более точным по модели, оцененной с 2014 г.

ста курса, темпов роста цен на нефть и приростов ставки. Всего мы рассматриваем 4 возможных полинома, и они могут отличаться у разных регрессоров, тем самым мы имеем 64 различные спецификации. Затем строится псевдовневыборочный наукаст на тестовой выборке (с переносом левой точки тестовой выборки в тренировочную на каждом шаге), и лучшая комбинация полиномов выбирается на основе минимизации среднеквадратической ошибки прогноза на тестовой выборке. Можно сказать, что мы тестируем качество псевдовневыборочного наукастинга в 64 версиях модели (11). Далее будут представлены результаты для лучшей с точки зрения качества наукастинга модели. Помимо наукастов в каждой точке также строятся и прогнозы на 1 и 2 месяца, но лучшая модель выбирается именно на основе минимизации квадрата ошибки наукастов.

Так как для наукастинга в момент времени t нужны значения всех высокочастотных регрессоров в момент времени t , а их значения недоступны, когда не все недели в текущем месяце нам известны, то приходится использовать их прогноз. Аналогично при прогнозировании на 1 и 2 месяца вперед нам необходимо прогнозировать регрессоры в правой части.

При тестировании рядов на стационарность тестом ADF для всех недельных рядов выбиралось ненулевое количество запаздывающих разностей на основе информационного критерия Шварца. Однако модели со включением лагов на основе критерия Шварца для их первых разностей дают менее качественный прогноз вне выборки. Прогноз на 1, 2 и 3 недели вперед на рассматриваемой тестовой выборке 2018M1–2019M11 на основе моделей с лагами оказывается менее точным согласно корню из среднеквадратических ошибок (Root Mean Square Error, RMSE), чем нулевой прогноз. В случае MIACR нулевой прогноз оказывается лучше только на 2 и 3 недели вперед, на первом шаге модель с лагами превосходит нулевой прогноз всего на 6%, согласно отношению RMSE. В связи с этим мы будем использовать нулевой прогноз для недельных регрессоров в первых разностях.

3.3. Оценка модели

Наилучшей оказывается модель в следующей спецификации: зависимость инфляции от номинального курса задается полиномом Log-Cauchy с двумя параметрами, используется 11 недельных лагов, начиная с первого в текущем месяце, ставка входит в уравнение с весами из полинома Nbeta с двумя параметрами и с использованием 6 лагов, цена на нефть входит с весами полинома Almon, и выбирается 4 недельных лага. Также в качестве регрессора был выбран 1 лаг инфляции:

$$\begin{aligned} \hat{\pi}_t = & \hat{\alpha}_0 + \hat{\alpha}_1^{\text{inf}} \pi_{t-1} + \\ & + \hat{\beta}^e * \left(\sum_{i=1}^4 w_{\log-C}^e (\hat{\theta}_1^e, \hat{\theta}_2^e) \Delta \ln e_{t,i} + \sum_{i=1}^4 w_{\log-C}^e (\hat{\theta}_1^e, \hat{\theta}_2^e) \Delta \ln e_{t-1,i} + \right. \\ & \left. + \sum_{i=1}^3 w_{\log-C}^e (\hat{\theta}_1^e, \hat{\theta}_2^e) \Delta \ln e_{t-2,i} \right) + \\ & + \hat{\beta}^m * \left(\sum_{i=1}^4 w_{Nbeta}^m (\hat{\theta}_1^m, \hat{\theta}_2^m) \Delta m_{t,i} + \sum_{i=1}^2 w_{Nbeta}^m (\hat{\theta}_1^m, \hat{\theta}_2^m) \Delta m_{t-1,i} \right) + \\ & + \sum_{i=1}^4 w_{Almon}^b (\hat{\theta}_1^b, \hat{\theta}_2^b) \Delta \ln b_{t,i}. \end{aligned} \quad (12)$$

Таблица 2. Результаты оценки параметров модели MIDAS

Параметр	Оценка	Стандартное отклонение	t-статистика	P-value
α_0	0,002	0,0003	6,038	0,0000 ^{***}
α_1^{inf}	0,591	0,035	16,853	0,0000 ^{***}
β^e	0,098	0,029	3,397	0,0010 ^{***}
θ_1^e	0,982	4,566	0,215	0,8300
θ_2^e	0,972	2,717	0,358	0,7210
β^m	0,006	0,001	7,809	0,0000 ^{***}
θ_1^m	1,096	0,155	7,074	0,0000 ^{***}
θ_2^m	0,988	0,004	235,197	0,0000 ^{***}
θ_1^b	-0,009	0,005	-1,770	0,0800 [*]
θ_2^b	0,003	0,002	1,711	0,0900 [*]

Примечание. Условные обозначения см. в примечании к табл. 1.

В случае Bridge-модели мы имеем дело с простой линейной регрессией на месячных данных, в рамках которой при построении наукастов будут использоваться известные за некоторый период значения (1–4 недели текущего месяца). На основе критерия Шварца выбрана спецификация Bridge-модели, в которой в качестве регрессоров выступают темп роста курса за прошлый месяц, темп роста цены на нефть за текущий месяц и изменение ставки за текущей месяц и за прошлый, а также 1 лаг инфляции.

Оценки модели (12) представлены в табл. 2.

Интерпретация коэффициентов в MIDAS-модели затруднительна в связи с тем, что мы регрессируем на некоторую нелинейную комбинацию наших высокочастотных регрессоров. Однако мы вполне можем проинтерпретировать знаки у коэффициентов β вне полиномов. Знак оценки коэффициента перед темпами роста курса ожидаем и логичен. Увеличение темпа роста номинального курса рублей за доллар, то есть ослабление курса, ускоряет инфляцию. У цен на нефть такого параметра нет из-за специфичного полинома Алмона, однако мы можем проинтерпретировать влияние каждого недельного лага. Так, повышение темпов роста цен на нефть в первые две недели месяца замедляет инфляцию, а в последующие две недели, напротив, ускоряет. Что касается знака оценки перед приростом MIA CR, то он оказывается противоречивым. Повышение ставки, то есть ужесточение ДКП, увеличивает инфляцию, что противоречит теории. Однако данный эффект объясняется фактическими данными и наиболее красноречиво проявляется, например, на периоде 2015–2017 гг., когда резко возросшая инфляция сопровождалась повышением процентных ставок, а затем снижение ставок происходило одновременно

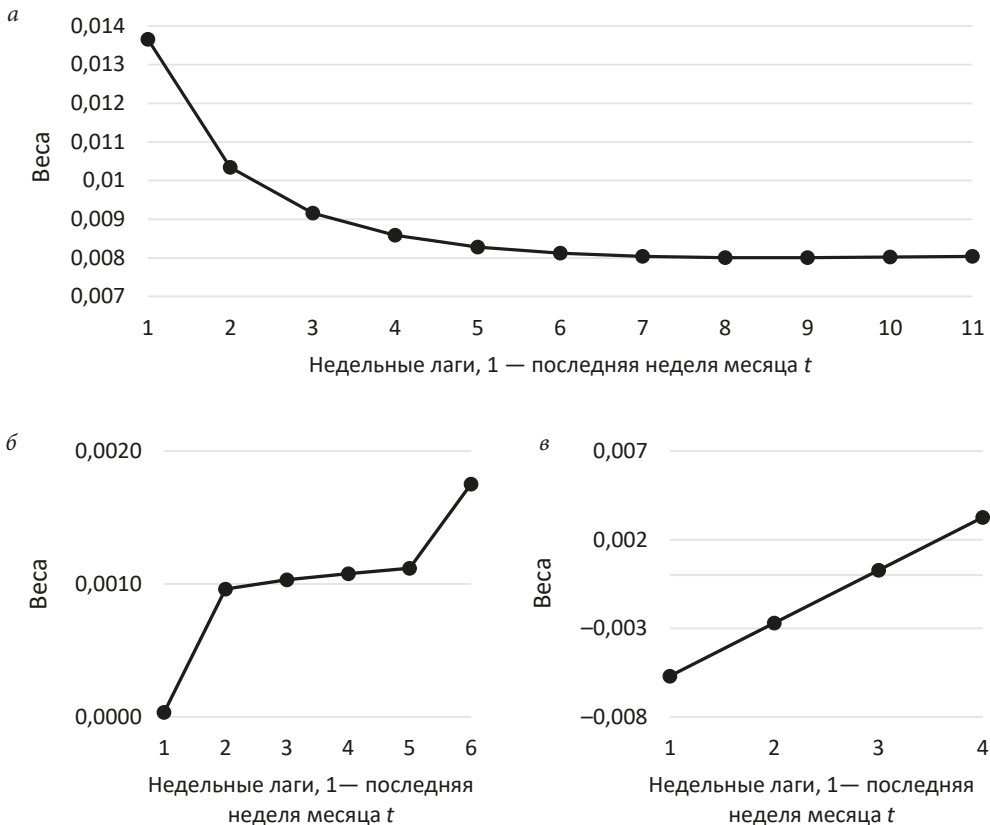


Рис. 2. Веса при недельных лагах: а — темпов роста номинального обменного курса; б — приростов ставки МІАСR; в — темпов роста цен на нефть; 1–4 — недели в текущем месяце t от последней до первой; 5–8 — недели от последней до первой в прошлом месяце $t - 1$; 9–11 — недели от последней до второй в позапрошлом месяце $t - 2$; t — текущий месяц

со снижением инфляции. Аналогичный результат был получен рядом авторов в работах по построению структурных моделей для анализа влияния шоков ДКП на цены [Пестова, 2018; Тишин, 2019].

Для наглядности на рис. 2 представлены полученные оценки весов при различных лагах для каждого из высокочастотных регрессоров.

4. Тестирование качества наукастинга и краткосрочного прогнозирования

Перейдем к сравнению качества прогнозов. Как было сказано ранее, мы рассматриваем четыре эксперимента при разной доступности недельных данных. В табл. 3 для формального сравнения представлены RMSE псевдодневных наукастов на выборке 2018M1–2019M11. Помимо моделей MIDAS и Bridge рассматриваются также три простых бенчмарка: 1) ARIMA с выбором лагов на основе

Таблица 3. RMSE моделей при наукастинге при разной доступности данных

Количество известных недель высокочастотных данных	MIDAS	Bridge	ARIMA	Последнее значение	Среднее за весь период
4 недели	0,001249	0,001477	0,001973	0,001520	0,003195
3 недели	0,001299	0,001476			
2 недели	0,001355	0,001513			
1 неделя	0,001407	0,001560			

Таблица 4. RMSE моделей при прогнозировании на 1 месяц при разной доступности данных

Количество известных недель высокочастотных данных	MIDAS	Bridge	ARIMA	Последнее значение	Среднее за весь период
4 недели	0,001699	0,001821	0,002580	0,001850	0,002991
3 недели	0,001700	0,001889			
2 недели	0,001659	0,001927			
1 неделя	0,001657	0,001977			

Таблица 5. RMSE моделей при прогнозировании на 2 месяца при разной доступности данных

Количество известных недель высокочастотных данных	MIDAS	Bridge	ARIMA	Последнее значение	Среднее за весь период
4 недели	0,001869	0,002065	0,002928	0,002104	0,002924
3 недели	0,001784	0,002113			
2 недели	0,001777	0,002195			
1 неделя	0,001855	0,002247			

критерия Шварца; 2) последнее значение инфляции (наивный прогноз); 3) среднее за период, на котором обучаются вышеописанные модели.

Как видно из таблицы, наилучший результат при любой доступности данных демонстрирует модель MIDAS. В случае, когда мы обладаем всеми данными (при четырех известных неделях), получается лучший прогноз. В модели Bridge аналогичная ситуация: ошибка при известных четырех или трех неделях отличается не сильно, а затем прогноз ухудшается.

Перейдем к рассмотрению прогнозов на 1 и 2 месяца. В табл. 4 и 5 представлены RMSE на тестовой выборке различных моделей при прогнозировании инфляции на 1 и 2 месяца вперед соответственно. В данных экспериментах известность данных за недели также обозначает известность в текущем месяце, все данные за недели в прошлых месяцах известны. Прогнозы строятся рекурсивно; наукаст используется в качестве значения инфляции в месяц t при построении прогноза на месяц $t + 1$. Аналогично при построении прогноза на месяц $t + 2$ в качестве значения лага инфляции используется прогноз, полученный на прошлом шаге (см. модель (12)).

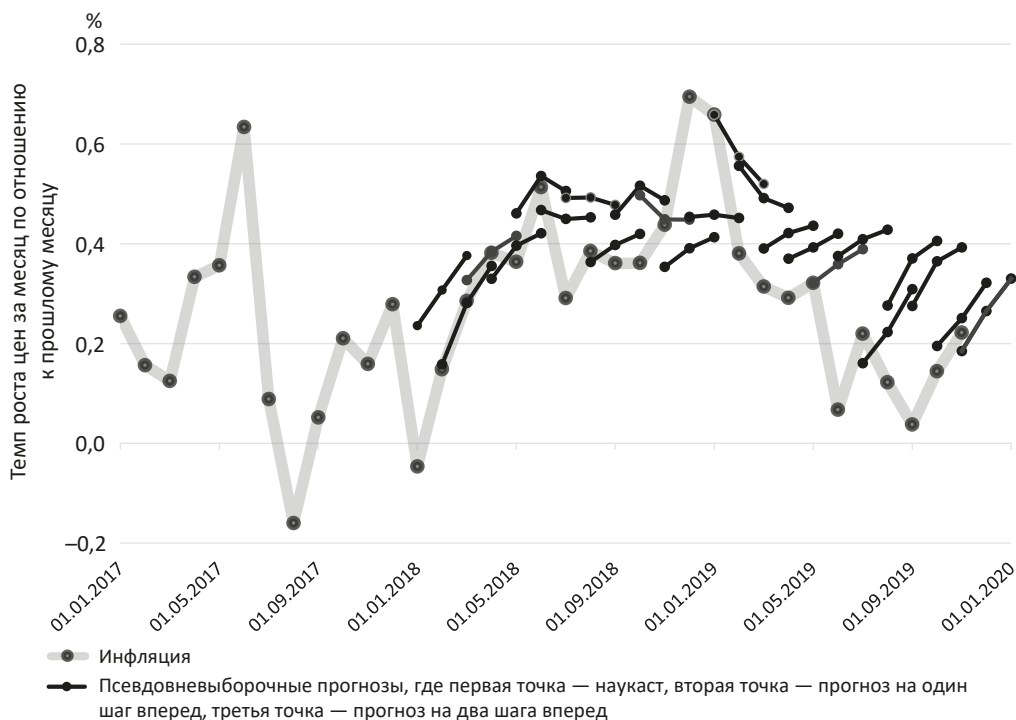


Рис. 3. Наукасты и прогнозы на 1 и 2 шага моделью MIDAS при известной 1 неделе высокочастотных данных

В обоих случаях MIDAS-модель также превосходит как Bridge, так и рассмотренные простые бенчмарки. Точность прогноза при увеличении горизонта ожидания падает: наукаст точнее прогноза на 1 и 2 месяца вперед, прогноз на 1 месяц вперед точнее, чем прогноз на 2 месяца вперед. В обоих случаях отсутствует тенденция уменьшения ошибки прогноза при увеличении числа известных недель в текущем месяце, что тоже ожидаемо, поскольку при увеличении горизонта прогноза известность данных в месяц t влияет слабее.

Для визуального представления на рис. 3 изображены псевдовневыборочные прогнозы, включающие наукаст, прогноз на 1 и 2 месяца по MIDAS-модели при известных высокочастотных регрессорах за 1 неделю (самый строгий эксперимент). Каждая линия из 3 точек представляет собой последовательно соединенные: наукаст, прогноз на 1 месяц и прогноз на 2 месяца. При построении графика используется самый строгий эксперимент, когда данные известны лишь за 1 неделю текущего месяца.

На рис. 2 можно заметить, что часто чем точнее мы строим наукаст, тем точнее прогноз на 1 месяц и на 2 месяца вперед. Что ожидаемо, так как прогноз строится рекурсивно. Из визуального анализа можно также заметить, что прогнозы стягиваются к слишком высокому среднему уровню. Особенно это заметно в 2019 г., когда инфляция существенно замедлилась относительно 2018 г. Этот факт объясняется периодом обучения модели 2010–2017 гг., на котором наблюдались значительно более высокие средние темпы инфляции, чем на периоде, на который строился

прогноз. Условно можно предположить наличие структурного сдвига в среднем уровне, который, по-видимому, произошел около 2016–2017 гг., когда режим таргетирования инфляции уже действовал и инфляция снизилась до рекордно низких темпов.

Мы оставляем задачу тестирования потенциальных структурных сдвигов для дальнейших исследований. В рамках данной работы с целью построения прогнозной модели предлагаем простой способ учета разных средних уровней в разные периоды времени, а именно включение в модель дамми-переменной, равной нулю до даты условного структурного сдвига и единице после. Дата будет выбрана на основе максимизации функции правдоподобия. Такой подход основан на известной особенности стационарных моделей, которая состоит в том, что математическое ожидание объясняемой переменной выражается через нелинейную комбинацию параметров модели (при нулевых математических ожиданиях экзогенных переменных). В рамках модели (11) математическое ожидание инфляции рассчитывается по простой формуле

$$\frac{\alpha_0}{1 - \alpha_1^{\text{inf}}}, \quad (13)$$

а в случае добавления в модель дамми-переменной, с коэффициентом λ перед ней, математическое ожидание после структурного сдвига рассчитывается по формуле

$$\frac{\alpha_0 + \lambda}{1 - \alpha_1^{\text{inf}}}. \quad (14)$$

Таким образом, мы ожидаем отрицательную оценку параметра λ , чтобы среднее после условного структурного сдвига стало ниже.

Датировку произведем с помощью следующей процедуры. Мы включаем дамми-переменную в модель, которая равна единице в различные периоды времени начиная с января 2014 г. по конец выборки и заканчивая периодом с июня 2017 г. по конец рассматриваемой выборки. Затем мы рассчитываем для каждой модели с дамми-переменной значение функции правдоподобия и сравниваем их между собой. Наибольшее значение функции получается в модели со включением дамми-переменной на первый месяц 2017 г. То есть в итоговую регрессию MIDAS была включена дамми-переменная, равная 1 с января 2017 г. до конца выборки.

В табл. 6 представлены результаты оценки модели с добавлением дамми-переменной.

Как видно из таблицы, $\hat{\lambda}$ значима на лишь 10%-м уровне. Поскольку условный структурный сдвиг предполагается достаточно близко к началу прогнозного периода, это ожидаемый результат, так как точек на оценивание данного параметра немного. Основной причиной добавления фиктивной переменной является повышение прогнозной силы модели, что видно из расчетов RMSE для новой модели, представленных в табл. 7.

Как видно из табл. 7, а также табл. 3–5, во всех случаях MIDAS-модель со включением фиктивной переменной превосходит по качеству аналогичную модель без ее добавления. Для наглядной визуализации прогнозов по расширенной модели обратимся к рис. 4, на котором изображена динамика инфляции месяц к прошлому

Таблица 6. Результаты оценки параметров модели MIDAS с дамми-переменной

Параметр	Оценка	Стандартное отклонение	t-статистика	P-value
α_0	0,002	0,0004	4,905	0,0000***
α_1^{inf}	0,557	0,041	13,628	0,000***
λ	-0,001	0,0005	-1,972	0,0510*
β^e	0,104	0,028	3,727	0,0000***
θ_1^e	0,967	4,564	0,212	0,8330
θ_2^e	0,950	2,799	0,339	0,7350
β^m	0,006	0,001	7,789	0,0000***
θ_1^m	1,109	0,188	5,913	0,0000***
θ_2^m	0,987	0,005	183,442	0,0000***
θ_1^b	-0,008	0,005	-1,616	0,1091
θ_2^b	0,003	0,002	1,660	0,0900*

Примечание. Условные обозначения см. в примечании к табл. 1.

Таблица 7. RMSE модели MIDAS с дамми-переменной при наукастинге и прогнозировании на 1 и 2 месяца при разной доступности данных

Количество известных недель высокочастотных данных	Наукаст	Прогноз на 1 месяц	Прогноз на 2 месяца
4 недели	0,001102	0,001345	0,001574
3 недели	0,001169	0,001404	0,001512
2 недели	0,001233	0,001381	0,001452
1 неделя	0,001339	0,001471	0,001668

месяцу, а также наукаста и прогноза на 1 и 2 месяца вперед. Дополнительно обратимся к рис. 5, на котором, помимо динамики инфляции, изображен 68 %-й доверительный интервал для наукастов.

Как видно из рис. 4, прогнозы в последние годы стали сходиться к намного более низкому среднему уровню и стали ближе к фактической инфляции, чем без добавления фиктивной переменной. Из рис. 5 мы можем заметить, что 68 %-й доверительный интервал накрывает фактический ряд инфляции почти во всех точках, за исключением нескольких пиковых точек.

Также заметим, что выше был построен прогноз для сезонно-очищенной инфляции за месяц по отношению к прошлому месяцу. Возможно, более привычным для анализа является показатель темпа роста ИПЦ за месяц к аналогичному месяцу предыдущего года, с учетом, что в динамике данного показателя отсутствует сезонность. На рис. 6 представлен график данного показателя и аналогично пересчитанные прогнозы.

Как видно на графике, наукасты оказываются предельно близки к фактической динамике, однако при анализе данного ряда скорее интересны прогнозы на более длинные горизонты. В целом прогноз на 2 месяца вперед достаточно близок к наб-

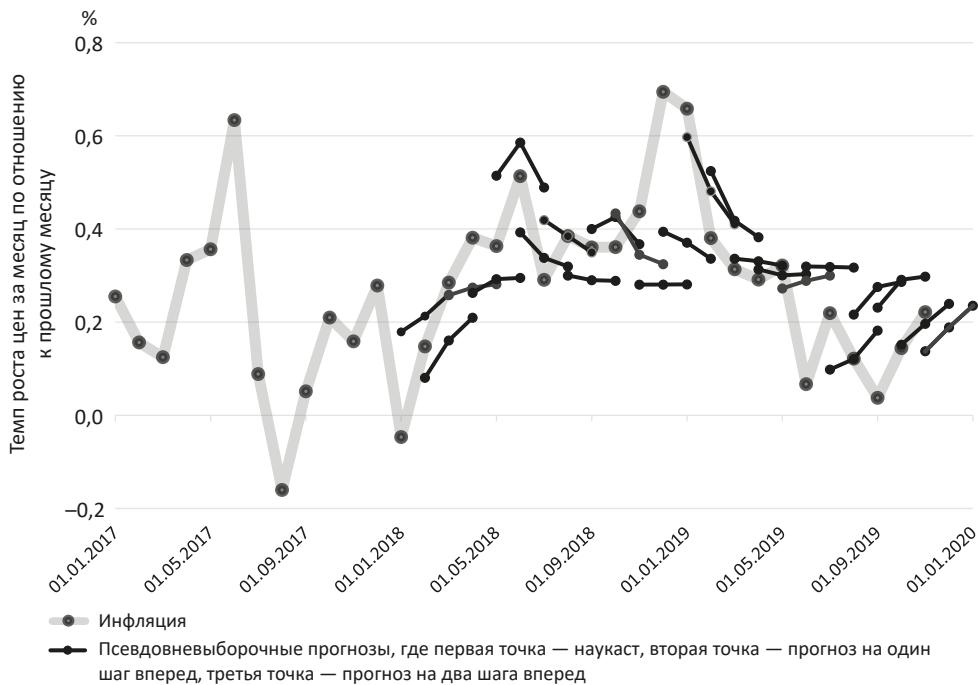


Рис. 4. Наукасты и прогнозы на 1 и 2 шага при известной 1 неделе высокочастотных данных, модель с добавлением дамми-переменной

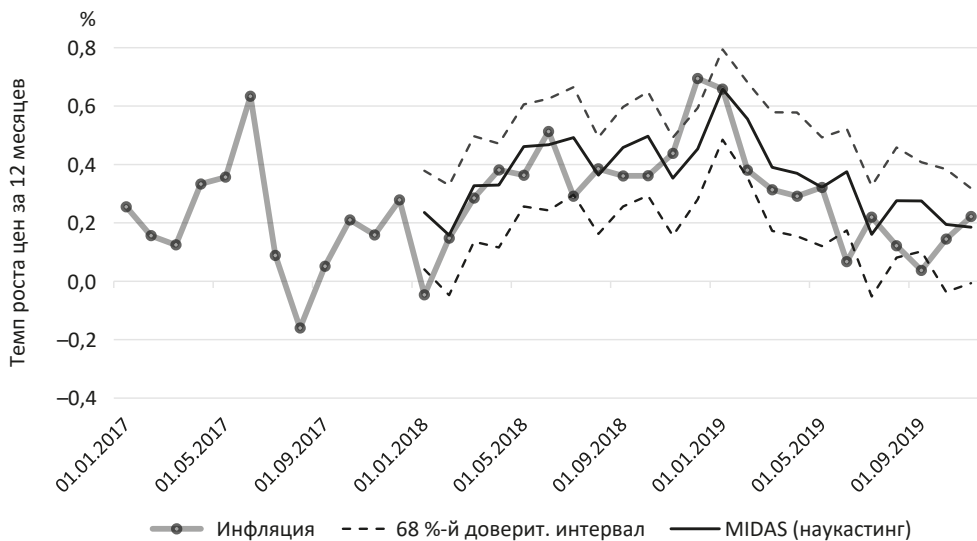


Рис. 5. Наукасты при известной 1 неделе высокочастотных данных и 68%-м доверительном интервале наукастов, модель с добавлением дамми-переменной

людаемому ряду и следует его динамике, за исключением нескольких периодов. Так, например, максимальная ошибка прогноза произошла в феврале 2019 г. — 0,9 п. п.

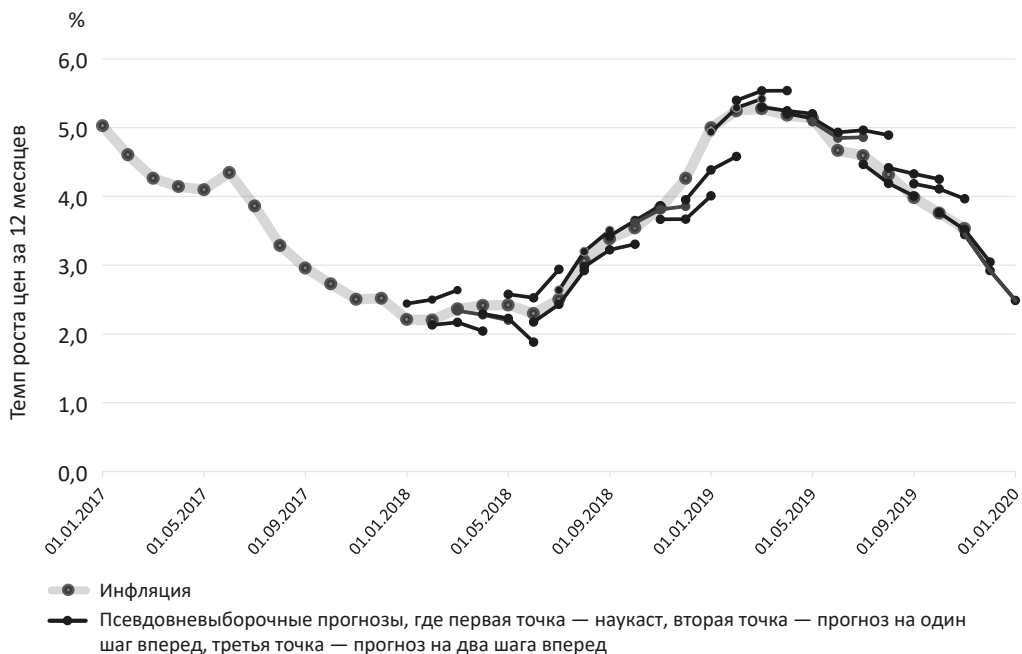


Рис. 6. Наукасты и прогнозы на 1 и 2 шага инфляции за 12 месяцев при известной 1 неделе высокочастотных данных, модель с добавлением дамми-переменной

Заключение

С учетом перехода к режиму таргетирования инфляции в 2014 г. задача построения и публикации качественных прогнозов стала критически важной для Центрального банка. В данной работе была апробирована модель MIDAS для наукастинга и прогнозирования месячной российской инфляции с помощью недельных данных по номинальному обменному курсу, ставке МІАСR и ценам на нефть. Рассматриваемая модель продемонстрировала высокую прогнозную силу относительно простых эконометрических бенчмарков, и в том числе относительно модели Bridge, также предназначенной для связи низкочастотных данных с высокочастотными.

С учетом небольшого пласта отечественной литературы на тему прогнозирования инфляции результаты работы, на наш взгляд, представляются весьма актуальными и могут быть использованы при усовершенствовании модельного аппарата для прогнозирования инфляции Банком России.

Благодарность

Авторы выражают благодарность анонимному рецензенту и А. В. Полбину за ценные комментарии к данной работе.

Литература

- Андреев А. (2016) Прогнозирование инфляции методом комбинирования прогнозов в Банке России. *Банк России. Серия докладов об экономических исследованиях*. Т. 14. С. 2–11.
- Байбуза И. (2018) Прогнозирование инфляции с помощью методов машинного обучения. *Деньги и кредит*. Т. 77, № 4. С. 42–59.
- Гафаров Б. Н. (2011) Кривая Филлипса и становление рынка труда в России. *Экономический журнал Высшей школы экономики*. Т. 15, № 2. С. 155–176.
- Зубарев А. В. (2018) Об оценке кривой Филлипса для российской экономики. *Экономический журнал Высшей школы экономики*. Т. 22, № 1. С. 40–58.
- Микош Х., Соланко Л. (2019) Прогнозирование роста российского ВВП с использованием данных со смешанной периодичностью. *Деньги и кредит*. Т. 78, № 1. С. 19–35.
- Павлов Е. (2020) Прогнозирование инфляции в России с помощью нейронных сетей. *Деньги и кредит*. Т. 79, № 1. С. 57–73.
- Перевычайин Ю., Перевычайина Е. (2015) Эффект переноса процентных ставок в России в 2010–2014 годах. *Экономическая политика*. Т. 10, № 5. С. 38–52.
- Пестова А. А. (2018) Об оценке эффектов монетарной политики в России: роль пространства шоков и изменений режимов политики. *Вопросы экономики*. № 2. С. 33–55.
- Пономарев Ю., Трунин П., Улюкаев А. (2014) Эффект переноса динамики обменного курса на цены в России. *Вопросы экономики*. № 3. С. 21–35.
- Стырин К. (2019) Прогнозирование инфляции в России методом динамического усреднения моделей. *Деньги и кредит*. Т. 78, № 1. С. 3–18.
- Тишин А. (2019) Неожиданные шоки ДКП в России. *Деньги и кредит*. Т. 78, № 4. С. 48–70.
- Angelini E., Bańbura M., Rünstler G. (2010). Estimating and forecasting the euro area monthly national accounts from a dynamic factor model. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, no. 1, pp. 1–22.
- Baffigi A., Golinelli R., Parigi G. (2004). Bridge models to forecast the euro area GDP. *International Journal of Forecasting*, vol. 20, no. 3, pp. 447–460.
- Baumeister C., Guérin P., Kilian L. (2015) Do high-frequency financial data help forecast oil prices? The MIDAS touch at work. *International Journal of Forecasting*, vol. 31, no. 2, pp. 238–252.
- Breitung J., Roling C. (2015) Forecasting inflation rates using daily data: A nonparametric MIDAS approach. *Journal of Forecasting*, vol. 34, no. 7, pp. 588–603.
- Delle Monache D., Petrella I. (2017) Adaptive models and heavy tails with an application to inflation forecasting. *International Journal of Forecasting*, vol. 33, no. 2, pp. 482–501.
- Galí J., Gertler M. (1999) Inflation dynamics: A structural econometric analysis. *Journal of monetary Economics*, vol. 44, no. 2, pp. 195–222.
- Ghysels E., Santa-Clara P., Valkanov R. (2004) The MIDAS touch: Mixed data sampling regression models. *Série Scientifique*, Mai. Montréal. 34 p.
- Ghysels E., Sinko A., Valkanov R. (2007) MIDAS regressions: Further results and new directions. *Econometric Reviews*, vol. 26, no. 1, pp. 53–90.
- Ghysels E., Kvedaras V., Zemlys V. (2016) Mixed frequency data sampling regression models: the R package midasr. *Journal of statistical software*, vol. 72, no. 4, pp. 1–35.
- Lucas R. E. (1973) Some international evidence on output-inflation tradeoffs. *The American Economic Review*, vol. 63, no. 3, pp. 326–334.
- McKnight S., Mihailov A., Rumler F. (2020) Inflation forecasting using the New Keynesian Phillips Curve with a time-varying trend. *Economic Modelling*, vol. 87, pp. 383–393.
- 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*, vol. 39, no. 1, pp. 98–119.
- Parigi G., Schlitzer G. (1995) Quarterly forecasts of the Italian business cycle by means of monthly economic indicators. *Journal of Forecasting*, vol. 14, no. 2, pp. 117–141.
- Ribon S., Suhoy T. (2011) Forecasting short run inflation using MIDAS. *Research Department, Bank of Israel*. 31 p.
- Sargent T. J., Wallace N. (1975) “Rational” Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule. *Journal of Political Economy*, vol. 83, no. 2, pp. 241–254.
- Sargent T. J., Wallace N. (1976) Rational expectations and the theory of economic policy. *Journal of Monetary Economics*, vol. 2, no. 2, pp. 169–183.

- Stock J. H., Watson M. W. (2007) Why has US inflation become harder to forecast? *Journal of Money, Credit and banking*, vol. 39, pp. 3–33.
- Trehan B. (1989) Forecasting growth in current quarter real GNP. *Economic Review-Federal Reserve Bank of San Francisco*, no. 1, p. 39.

Статья поступила в редакцию: 22.08.2020
Статья рекомендована в печать: 30.03.2021

Контактная информация:

Третьяков Дмитрий Витальевич — студент, мл. науч. сотр.; tretyakov-dv@ranepa.ru
Фокин Никита Денисович — аспирант, науч. сотр.; fokinikita@gmail.com

Does the high-frequency data is helpful for forecasting Russian inflation?*

D. V. Tretyakov, N. D. Fokin

Russian Academy of National Economy and Public Administration,
82, pr. Vernadskogo, Moscow, 117517, Russian Federation

For citation: Tretyakov D. V., Fokin N. D. (2021) Does the high-frequency data is helpful for forecasting Russian inflation? *St Petersburg University Journal of Economic Studies*, vol. 37, iss. 2, pp. 318–343. <https://doi.org/10.21638/spbu05.2021.206>

Due to the fact that at the end of 2014 the Central Bank made the transition to a new monetary policy regime for Russia — the inflation targeting regime, the problem of forecasting inflation rates became more relevant than ever. In the new monetary policy regime, it is important for the Bank of Russia to estimate the future inflation rate as quickly as possible in order to take measures to return inflation to the target level. In addition, for effective monetary policy, the households must trust the actions of monetary authorities and they must be aware of the future dynamics of inflation. Thus, to manage inflationary expectations of economic agents, the Central Bank should actively use the information channel, publish accurate forecasts of consumer price growth. The aim of this work is to build a model for nowcasting, as well as short-term forecasting of the rate of Russian inflation using high-frequency data. Using this type of data in models for forecasting is very promising, since this approach allows to use more information about the dynamics of macroeconomic indicators. The paper shows that using MIDAS model with weekly frequency series (RUB/USD exchange rate, the interbank rate MIACR, oil prices) has more accurate forecast of monthly inflation compared to several basic models, which only use low-frequency data.

Keywords: inflation, nowcasting, forecasting, high-frequency data, MIDAS model.

References

- Andreev A. (2016) Forecasting inflation by combining forecasts at the Bank of Russia. Bank of Russia. *Economic Research Report Series*, vol. 14, pp. 2–11. (In Russian)
- Angelini E., Bańbura M., Rünstler G. (2010) Estimating and forecasting the euro area monthly national accounts from a dynamic factor model. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, no. 1, pp. 1–22.
- Baffigi A., Golinelli R., Parigi G. (2004). Bridge models to forecast the euro area GDP. *International Journal of forecasting*, vol. 20, no. 3, pp. 447–460.
- Baumeister C., Guérin P., Kilian L. (2015) Do high-frequency financial data help forecast oil prices? The MIDAS touch at work. *International Journal of Forecasting*, vol. 31, no. 2, pp. 238–252.
- Baybuza I. (2018) Inflation Forecasting Using Machine Learning Methods. *Russian Journal of Money and Finance*, vol. 77, no. 4, pp. 42–59. (In Russian)

* The article was written on the basis of the Russian Academy of National Economy and Public Administration state assignment research programme.

- Breitung J., Røling C. (2015) Forecasting inflation rates using daily data: A nonparametric MIDAS approach. *Journal of Forecasting*, vol. 34, no. 7, pp. 588–603.
- Delle Monache, D., Petrella, I. (2017). Adaptive models and heavy tails with an application to inflation forecasting. *International Journal of Forecasting*, vol. 33, no. 2, pp. 482–501.
- Gali J., Gertler M. (1999) Inflation dynamics: A structural econometric analysis. *Journal of monetary Economics*, vol. 44, no. 2, pp. 195–222.
- Gafarov B. N. (2011) Phillips curve and the formation of the labor market in Russia. *Economic Journal of the Higher School of Economics*, vol. 15, no. 2, pp. 155–176. (In Russian)
- Ghysels E., Santa-Clara P., Valkanov R. (2004) The MIDAS touch: Mixed data sampling regression models. *Série Scientifique*, Mai. Montréal. 34 p.
- Ghysels E., Sinko A., Valkanov R. (2007) MIDAS regressions: Further results and new directions. *Econometric Reviews*, vol. 26, no. 1, pp. 53–90.
- Ghysels E., Kvedaras V., Zemlys V. (2016) Mixed frequency data sampling regression models: the R package midasr. *Journal of statistical software*, vol. 72, no. 4, pp. 1–35.
- Lucas R.E. (1973) Some international evidence on output-inflation tradeoffs. *The American Economic Review*, vol. 63, no. 3, pp. 326–334.
- McKnight S., Mihailov A., Rumler F. (2020) Inflation forecasting using the New Keynesian Phillips Curve with a time-varying trend. *Economic Modelling*, vol. 87, pp. 383–393.
- 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*, vol. 39, no. 1, pp. 98–119.
- Mikosch H., Solanko L. (2019) Forecasting Quarterly Russian GDP Growth with Mixed-Frequency Data. *Russian Journal of Money and Finance*, vol. 78, no. 1, pp. 19–35. (In Russian)
- Pavlov E. (2020) Forecasting Inflation in Russia Using Neural Networks. *Russian Journal of Money and Finance*, vol. 79, no. 1, pp. 57–73. (In Russian)
- Parigi G., Schlitzer G. (1995) Quarterly forecasts of the Italian business cycle by means of monthly economic indicators. *Journal of Forecasting*, vol. 14, no. 2, pp. 117–141.
- Perevyshin Yu., Perevyshina E. (2015) The effect of the transfer of interest rates in Russia in 2010–2014. *Economic Policy*, vol. 10, no. 5, pp. 38–52. (In Russian)
- Pestova A. A. (2018) On the effects of monetary policy in Russia: The role of the space of spanned and the policy regime shifts. *Voprosy Ekonomiki*, no. 2, pp. 33–55. (In Russian)
- Ponomarev Y., Trunin P., Ulyukaev A. (2014) Exchange Rate Pass-through in Russia. *Voprosy Ekonomiki*, no. 3, pp. 21–55. (In Russian)
- Ribon S., Suhoy T. (2011) Forecasting short run inflation using MIDAS. *Research Department, Bank of Israel*. 31 p.
- Sargent T. J., Wallace N. (1975) “Rational” Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule. *Journal of Political Economy*, vol. 83, no. 2, pp. 241–254.
- Sargent T. J., Wallace N. (1976) Rational expectations and the theory of economic policy. *Journal of Monetary Economics*, vol. 2, no. 2, pp. 169–183.
- Stock J. H., Watson M. W. (2007) Why has US inflation become harder to forecast? *Journal of Money, Credit and banking*, vol. 39, pp. 3–33.
- Styrin K. (2019) Forecasting Inflation in Russia Using Dynamic Model Averaging. *Russian Journal of Money and Finance*, vol. 78, no. 1, pp. 3–18. (In Russian)
- Tishin A. (2019) Monetary Policy Surprises in Russia. *Russian Journal of Money and Finance*, vol. 78, no. 4, pp. 48–70. (In Russian)
- Trehan B. (1989) Forecasting growth in current quarter real GNP. *Economic Review-Federal Reserve Bank of San Francisco*, no. 1, p. 39.
- Zubarev A. (2018) On the Estimation of the Phillips Curve for the Russian Economy. *HSE Economic Journal*, vol. 22, no. 1, pp. 40–58. (In Russian)

Received: 22.08.2020

Accepted: 30.03.2021

Authors' information:

Dmitriy V. Tretyakov — Student, Junior Research Fellow; tretyakov-dv@ranepa.ru

Nikita D. Fokin — Postgraduate Student, Research Fellow; fokinikita@gmail.com