

Научная статья

УДК 338.2

JEL E32, E37, E52, E70

DOI 10.25205/2542-0429-2023-23-4-60-82

Учет информационного фона в DSGE-модели экономики России с адаптивным обучением¹

Дмитрий Васильевич Колюжнов¹

Егор Дмитриевич Колюжнов²

Маргарита Валерьевна Ляхнова³

^{1,3}Институт экономики и организации промышленного производства СО РАН
Новосибирск, Россия

^{1,2}Новосибирский национальный исследовательский государственный университет
Новосибирск, Россия

³Банк России
Москва, Россия

¹dima.kolyuzhnov@mail.ru, <https://orcid.org/0000-0003-4390-7289>

²e.kolyuzhnov@g.nsu.ru

³rita_2000@list.ru, <https://orcid.org/0000-0003-1057-1387>

Аннотация

В рамках данного исследования разрабатывается метод применения результатов модели анализа информационного фона в модели формирования ожиданий адаптивно обучающихся экономических агентов в общей постановке DSGE-модели. Этот метод тестируется на малой DSGE-модели экономики России с адаптивным обучением, разработанной нами в ИЭОПП СО РАН. На примере этой модели показывается, что предлагаемый метод улучшает соответствие данных, имитируемых моделью, экономической статистике, что позволяет использовать эту модель для прогнозирования макроэкономических показателей, рассматривая различные сценарии развития экономики при разной окраске будущего информационного фона. Делается вывод, что управление новостным потоком оказывает влияние на функционирование экономики и может потенциально использоваться как элемент экономической политики, последствия которой можно оценить, используя наш метод. Универсальность метода, предложенного в работе, позволяет распространить его применение на широкий ряд DSGE-моделей, используемых центральными банками большинства стран мира.

¹ Настоящая статья отражает личную позицию авторов. Содержание и результаты данного исследования не следует рассматривать, в том числе цитировать в каких-либо изданиях, как официальную позицию Банка России или указание на официальную политику или решения регулятора. Любые ошибки в данном материале являются исключительно авторскими.

Ключевые слова

DSGE-моделирование, адаптивное обучение, анализ тональности, нейронные сети, машинное обучение, российская экономика, прогнозы

Источник финансирования

Работа выполнена по плану НИР ИЭОПП СО РАН, проект «Методы и модели обоснования стратегии развития экономики России в условиях меняющейся макроэкономической реальности», № 121040100281-8

Для цитирования

Колюжнов Д. В., Колюжнов Е. Д., Ляхнова М. В. Учет информационного фона в DSGE-модели экономики России с адаптивным обучением // Мир экономики и управления. 2023. Т. 23, № 4. С. 60–82. DOI 10.25205/2542-0429-2023-23-4-60-82

Including the Information Background into the DSGE Model of the Russian Economy with Adaptive Learning

Dmitri V. Kolyuzhnov¹, Egor D. Kolyuzhnov²
Margarita V. Lyakhnova³

^{1,3}Institute of Economics and Industrial Engineering SB RAS
Novosibirsk, Russian Federation

^{1,2}Novosibirsk State University,
Novosibirsk, Russian Federation

³Bank of Russia,
Moscow, Russian Federation

¹dima.kolyuzhnov@mail.ru, <https://orcid.org/0000-0003-4390-7289>

²e.kolyuzhnov@g.nsu.ru

³rita_2000@list.ru, <https://orcid.org/0000-0003-1057-1387>

Abstract

Within the framework of this study, a method is being developed for applying the results of the information background analysis model to the expectations formation model of adaptively learning economic agents in the general formulation of the DSGE model. This method is tested on a small DSGE model of the Russian economy with adaptive learning, developed by us at the IEIE SB RAS. Using this model as an example, we show that the proposed method improves the fit of the data simulated by the model to economic statistics, which makes it possible to use this model to predict macroeconomic indicators, comparing different scenarios for economic development depending on the future information background sentiment. We conclude that news flow management has an impact on the economy performance and can potentially be used as an element of economic policy, whose consequences can be evaluated using our method. The versatility of the method proposed in this paper allows its application to be extended to a wide range of DSGE models used by central banks in most countries of the world.

Keywords

DSGE-modeling, adaptive learning, sentiment analysis, neural networks, machine learning, Russian economy, forecasts

Funding

The publication is prepared within the project “Methods and models of Russian economy development strategy justification under changing macroeconomic reality”, № 0260-2021-0008 according to the research plan of the IEIE SB RAS.

For citation

Kolyuzhnov D. V., Kolyuzhnov E. D., Lyakhnova M. V. Including the information background into the DSGE model of the Russian economy with adaptive learning. *World of Economics and Management*, 2023, vol. 23, no. 4, pp. 60–82. (in Russ.) DOI 10.25205/2542-0429-2023-23-4-60-82

Введение

Влияние ожиданий на принятие решений и результирующую динамику показателей, характеризующих поведение системы, – это то, что отличает экономику от других наук. Учет ожиданий экономических агентов и правильное описание моделей их формирования является необходимым условием для написания моделей, адекватных экономической реальности, которые можно использовать для отработки последствий принятия различных решений в области экономической политики и для прогнозирования. В свою очередь, не секрет, что в современном мире ожидания формируются в том числе под воздействием информационного фона, формируемого новостями, поступающими из различных средств массовой информации и дублируемыми в соцсетях. Таким образом, учет влияния новостей посредством информационного фона на формирование ожиданий является перспективным направлением в экономическом моделировании, в частности, в построении макроэкономических моделей.

Современные методы позволяют количественно учесть влияние информационного фона на формирование ожиданий через преобразование потока новостной информации в численные метрики посредством *машинного обучения*. Соответствующую данному информационному потоку метрику можно таким образом использовать в различных экономических моделях, в том числе в моделях формирования ожиданий.

Естественной областью применения, где необходимо учитывать влияние информационного фона на формирование ожиданий и, соответственно, на результирующую экономическую динамику, являются широко распространенные модели, используемые регуляторами по всему миру, а именно динамические стохастические модели общего экономического равновесия (DSGE-модели). В настоящее время DSGE-модели широко применяются не только в академической среде, но и крупнейшими центральными банками и международными финансовыми организациями. Примерами построенных DSGE-моделей являются модель NAWM (ЕЦБ) [1; 2], модель EDO (ФРС США) [3; 4], модели ToTEM и ToTEM II (Банк Канады) [5; 6], модели BEQM и COMPASS (Банк Англии) [7; 8], модель NEMO (Норвежский Банк) [9], модели GEM и GIMF (МВФ) [10], большая и малая DSGE-модели (Банк России) [11; 12], и множество других DSGE-моделей, разработанных и используемых центральными банками по всему миру.

Большинство используемых регуляторами DSGE-моделей находятся в рамках «мейнстрима» при моделировании формирования ожиданий экономическими агентами, а именно задействуют гипотезу о рациональных ожиданиях (далее RE). Эта гипотеза является очень сильной, предполагающей знание агентами истинной модели (и ее параметров), лежащей в основе экономической динамики, распределений всех шоков, и предполагает верное представление о поведении других агентов и координацию действий. В то же время в эмпирических исследованиях

экономисты, предполагающие равновесия при рациональных ожиданиях в своих теоретических моделях, не знают значений параметров и вынуждены оценивать их эконометрически. Как обосновывает Сарджент [13], представляется более естественным предполагать при моделировании, что экономические агенты сталкиваются с такими же ограничениями, и рассматривать их как эконометристов (или статистиков) при прогнозировании будущего состояния экономики, которые оценивают прогнозные модели, используя стандартные статистические процедуры, такие как метод наименьших квадратов, метод стохастического градиента, байесовское обучение, и формируют убеждения о параметрах модели. Основанная на таких методах процедура пересмотра и обновления убеждений называется адаптивным (эконометрическим) обучением. Хороший обзор литературы, посвященной этой форме адаптивного обучения, представлен Эвансом и Хонкапойей [14]. Примером применения адаптивного обучения в DSGE-моделях является разработанная нами «Малая DSGE-модель экономики России с неоднородным адаптивным обучением» [15].

Целью данного исследования является разработка метода применения результатов модели анализа информационного фона в модели формирования ожиданий адаптивно обучающихся экономических агентов в общей постановке DSGE-модели и его последующее тестирование на разработанной нами DSGE-модели экономики России.

Подобных публикаций, где при адаптивном обучении экономических агентов в DSGE-модели, построенной для экономики России, учитывался бы информационный фон, обнаружено не было, поэтому в определенном смысле мы являемся первопроходцами в этом направлении.

Методика учета информационного фона в DSGE-моделях с адаптивным обучением

Предлагаемый в данной работе подход к учету информационного фона в DSGE-моделях с адаптивным обучением может быть применен к большинству существующих и перспективных линейных (или линеаризованных) экономических моделей с ожиданиями, записанных в сокращенной форме, относящихся, как у Богомоловой и Колужнова [16], к общему классу структурно-неоднородных линейных моделей с S -типами агентов с различными прогнозами, представленному в виде

$$y_t = \alpha + \sum_{i=1}^d L_i y_{t-i} + \sum_{h=1}^S \sum_{b=0}^m \sum_{f=b}^n A_{bf}^h \hat{E}_{t-b}^h y_{t-b+f} + B w_t + \zeta \varepsilon_t, A_{00}^h \equiv 0 \quad (1)$$

$$w_t = F w_{t-1} + v_t, \quad (2)$$

где y_t есть $n \times 1$ – вектор эндогенных переменных, $w_t - k \times 1$ – вектор экзогенных переменных; v_t и ε_t – векторы (независимых) шоков белого шума; $\hat{E}_{t-b}^h y_{t-b+f}$ есть (в общем случае нерациональные) ожидания эндогенных переменных агентом h ; а L_i, A_{bf}^h, B, ζ – подходящего размера матрицы. Структура F ($k \times k$ – матрица)

предполагает, что w_t следует VAR (1) стационарному процессу с матрицей ковариаций $M_w = \lim_{t \rightarrow \infty} E w_t w_t'$.

Матрицы A_{bf}^h включают в себя массу ζ_h каждого типа агента h , $\sum \zeta_h = 1$ и определены как $A_{bf}^h = \zeta_h \cdot \tilde{A}_{bf}^h$, где $\tilde{A}_{bf}^h$ показывают, как агенты типа h реагируют на свои собственные прогнозы и содержат структурные параметры, характеризующие данную экономику. В общем случае, A_{bf}^h могут быть разными для различных типов агентов, что представляет собой структурную неоднородность модели. Таким образом, в моделях из этого класса ожидания эндогенных переменных, формируемые различными типами агентов, линейно влияют на текущие значения этих переменных.

Наиболее часто используемыми в литературе примерами моделей из этого класса являются Модель I (с параметрами $d = 0$, $m = 1$, n – любое, $\tilde{A}_{bf}^h$), Модель II (с $A_{0f}^h \equiv 0$), Модель III (с $d = 0$, $m = 0$, n – любое) и Модель IV (с $d = 0$, $m = 1$, $n = 1$), задаваемые соответственно как (2) и

$$y_t = \alpha + \sum_{h=1}^S A_0^h \hat{E}_{t-1}^h y_t + \sum_{h=1}^S A_1^h \hat{E}_{t-1}^h y_{t+1} + \dots + \sum_{h=1}^S A_\tau^h \hat{E}_{t-1}^h y_{t+\tau} + B w_t + \zeta \varepsilon_t \text{ для Модели I,} \quad (3)$$

$$y_t = \alpha + L y_{t-1} + \sum_{h=1}^S A_0^h \hat{E}_{t-1}^h y_t + \sum_{h=1}^S A_1^h \hat{E}_{t-1}^h y_{t+1} + B w_t + \zeta \varepsilon_t \text{ для Модели II,} \quad (4)$$

$$y_t = \alpha + \sum_{h=1}^S A_1^h \hat{E}_t^h y_{t+1} + \dots + \sum_{h=1}^S A_\tau^h \hat{E}_t^h y_{t+\tau} + B w_t + \zeta \varepsilon_t \text{ для Модели III,} \quad (5)$$

$$y_t = \alpha + L y_{t-1} + \sum_{h=1}^S A_1^h \hat{E}_t^h y_{t+1} + B w_t + \zeta \varepsilon_t,$$

с информационным множеством $(1, y'_{t-1}, w'_t)$ для Модели IV, (6)

в которых ожидания агентов каждого типа h строятся, основываясь на воспринимаемом ими законе движения (PLM): $y_t = \Phi'_{h,t-1} w_{t-1}$ (для Моделей I и II) и $y_t = \Phi'_{h,t} w_t$ (для Моделей III и IV), где $\Phi'_{h,t-1} = (a_{h,t-1}, b_{h,t-1})$ для Модели I, $\Phi'_{h,t-1} = (a_{h,t-1}, b_{h,t-1}, c_{h,t-1})$ для Модели II, $\Phi'_{h,t-1} = (a_{h,t}, b_{h,t})$ для Модели III и $\Phi'_{h,t} = (a_{h,t}, b_{h,t}, c_{h,t})$ для Модели IV носят название параметров *убеждений*. В качестве отправной точки моделирования обычно рассматриваются PLM, соответствующие по структуре равновесию при рациональных ожиданиях (REE) с минимальным числом переменных состояния, однако на практике моделирования часто удобно предполагать, что агенты неверно специфицируют модель ввиду отсутствия у них полного представления об экономике или в силу ограниченности данных.

В общем случае в рассмотренных выше классах моделей предполагается, что агенты при формировании и обновлении ожиданий используют следующее неоднородное смешанное RLS/SG обучение (RLS – рекурсивный метод наименьших квадратов, используется агентами $h = 1, \dots, S_0$, SG – обучение по методу стохастического градиента, используется агентами $h = S_0 + 1, \dots, S$):

$$\begin{aligned} \Phi_{h,t+1} &= \Phi_{h,t} + \alpha_{h,t+1} R_{h,t+1}^{-1} z_t (y_{t+1} - \Phi'_{h,t} z_t)', \quad h = 1, \dots, S_0 \text{ (RLS)} \\ R_{h,t+1} &= R_{h,t} + \alpha_{h,t+1} (z_t z_t' - R_{h,t}) \end{aligned} \quad (7)$$

$$\Phi_{h,t+1} = \Phi_{h,t} + \alpha_{h,t+1} z_t' (y_{t+1} - \Phi_{h,t}' z_t'), h = S_0 + 1, \dots, S \text{ (SG)} \quad (8)$$

где $z_t' = (1, w_t')$ для Модели I, $z_t' = (1, y_t', w_t')$ для Модели II;

$$\begin{aligned} \Phi_{h,t+1} &= \Phi_{h,t} + \alpha_{h,t+1} R_{h,t+1}^{-1} z_t' (y_t - \Phi_{h,t}' z_t'), \\ R_{h,t+1} &= R_{h,t} + \alpha_{h,t+1} (z_t z_t' - R_{h,t}) \end{aligned} \quad h = 1, \dots, S_0 \text{ (RLS)} \quad (9)$$

$$\Phi_{h,t+1} = \Phi_{h,t} + \alpha_{h,t+1} z_t' (y_t - \Phi_{h,t}' z_t') \quad h = S_0 + 1, \dots, S \text{ (SG)} \quad (10)$$

где $z_t' = (1, w_t')$ для Модели III, $z_t' = (1, y_{t-1}', w_t')$ для Модели IV.

Неоднородность в обучении проявляется в виде разных типов алгоритмов обучения, используемых агентами (RLS и SG), различного начального восприятия (представленного начальными значениями для алгоритмов обучения для каждого агента) и разных долей агентов, использующих алгоритм обучения определенного типа. Разные скорости реагирования на новую информацию задаются разными степенями инерции $\delta_h > 0$ [17], постоянными коэффициентами перед детерминистической убывающей последовательностью коэффициентов приращения в алгоритме обучения α_t , общей для всех агентов, $\delta_h : \alpha_{h,t} = \delta_h \alpha_t, \forall h$ (детали см. в [16]). Случаи однородного обучения являются частными случаями общей постановки.

Изначально информационный фон никак не представлен ни в уравнениях, задающих структурную форму модели, ни в алгоритмах адаптивного обучения. Естественным способом включения такого фона в модель является предположение о влиянии информационного фона, выраженного в виде какой-то количественной метрики, на ожидания.

Обозначив полученную каким-то образом из новостного потока количественную метрику информационного фона для данной эндогенной переменной i в момент времени t через N_t^i , процесс формирования ожиданий агентами в рассмотренных выше классах с учетом информационного фона можно представить через модификацию их PLM следующим образом²:

$y_t = a_{h,t-1} + b_{h,t-1} w_{t-1} + d_{h,t-1} N_{t-1}$ (модифицированный PLM для Модели I), $y_t = a_{h,t-1} + b_{h,t-1} y_{t-1} + c_{h,t-1} w_{t-1} + d_{h,t-1} N_{t-1}$ (для Модели II), $y_t = a_{h,t} + b_{h,t} w_t + d_{h,t} N_{t-1}$ (для Модели III), $y_t = a_{h,t} + b_{h,t} y_{t-1} + c_{h,t} w_t + d_{h,t} N_{t-1}$ (для Модели IV), где N_t – вектор метрик информационного фона. В каждом конкретном случае моделирования процесса формирования ожиданий можно задавать определенную специфику включения тех или иных элементов вектора метрик информационного фона в процесс формирования ожиданий рассматриваемым типом агентов, например, путем зануления соответствующих коэффициентов в $d_{h,t-1}$ для всех периодов (т. е. путем априорного невключения отдельных факторов в PLM).

Соответственно, в алгоритмах смешанного RLS/SG обучения, представленных (7–10), произойдут следующие изменения в структуре векторов

² В данной постановке мы предполагаем, что информационный фон влияет на ожидания с лагом в один период; в определенных случаях, если это оправдано экономической средой, можно рассматривать и случай одновременного воздействия фона на предсказываемую переменную, модель позволяет это сделать.

убеждений и векторов факторов для учета включения информационного фона в модель формирования ожиданий: $\Phi'_{h,t-1} = (a_{h,t-1}, b_{h,t-1}, d_{h,t-1})$ и $z'_t = (1, w'_t, N'_t)$ для Модели I; $\Phi'_{h,t-1} = (a_{h,t-1}, b_{h,t-1}, c_{h,t-1}, d_{h,t-1})$ и $z'_t = (1, y'_t, w'_t, N'_t)$ для Модели II; $\Phi'_{h,t-1} = (a_{h,t}, b_{h,t}, d_{h,t})$ и $z'_t = (1, w'_t, N'_{t-1})$ для Модели III; и $\Phi'_{h,t} = (a_{h,t}, b_{h,t}, c_{h,t}, d_{h,t-1})$ и $z'_t = (1, y'_{t-1}, w'_t, N'_{t-1})$ для Модели IV.

Методика анализа информационного фона на основе потока новостей

Описание метода построения численных метрик информационного фона

Для построения численных метрик, соответствующих информационному фону, влияющему на формирование ожиданий эндогенных переменных в рассмотренных выше алгоритмах адаптивного обучения, необходимо провести предварительную подготовку текста новости – перевод его из текстовой формы в форму векторов, с которыми может работать метод *машинного обучения*. Конечная цель рассматриваемого подхода состоит в том, чтобы сопоставить потоку новостей за предшествующий период метрику – число (от -1 до 1), наилучшим образом отражающее влияние новости на прогнозируемый показатель. По сути, для этого требуется построение некоторой модели анализа информационного фона для каждой предсказываемой эндогенной переменной. В литературе известен подход к построению такого рода моделей, носящий название *анализ тональности* (Sentiment Analysis) – метод, позволяющий определить, носит ли содержание текста положительную, отрицательную или нейтральную окраску.

Для обработки текста новостей с целью дальнейшего использования их в алгоритмах *машинного обучения* применяется следующий конвейер: текст отдельной новости преобразовывается с помощью удаления ненужных слов и символов (знаков препинания и других символов, не являющихся буквами русского или английского алфавита), а затем оставшиеся слова в тексте преобразуются в основы слова (слова без окончаний). Последняя процедура называется *стемминг*. После этого применяется *векторизация* – преобразование слов в тексте новостей в вектор фиксированной длины.

В качестве методов *машинного обучения* для предсказания изменения макроэкономических показателей в данной работе используются *нейронные сети*. В качестве входных параметров для таких сетей служат векторы, полученные на предыдущем этапе обработки текстов новостей.

Нейронная сеть в общем виде представляет собой математическую модель, построенную по принципу организации и функционирования биологических нейронных сетей – функцию, преобразующую один вектор в другой. Самой простой нейронной сетью является функция вида $f(v) = Ya(Xv + b) + c$, где v – входной вектор размера m , X – матрица размера $k \times m$, a – функция активации (любая функция $a: R \rightarrow R$, которая выполняется векторно – $a([v_1 \dots v_m]) = [a(v_1) \dots a(v_m)]$, (часто в виде такой функции активации используют $ReLU(x) = \max(0, x)$ – Rectified Linear Unit), Y – матрица размера $n \times k$, где n – размер выходного вектора, который нам нужно получить, b и c – векторы смещений размеров k и n соответственно.

Одним из видов нейронных сетей, используемых в этой работе, является *полносвязная нейронная сеть* – сеть, состоящая из последовательного применения функций вида $a(Xv)$ – слоев. Когда говорят, что используется полносвязный слой размера $n \times m$, подразумевают, что используется матрица этого размера. Нейронные сети могут складываться из разных слоев (не обязательно описанного ранее типа) в одну цельную нейронную сеть. Результаты (вектора с последнего слоя) нейронных сетей оцениваются с помощью функции потерь – функции от полученных данных с помощью нейронной сети и от реальных данных. В данной работе для этих целей используется MSE (среднеквадратическая ошибка).

Каждая нейронная сеть может обучаться – обновлять содержимое матриц и векторов смещений – параметры по некоторому алгоритму. В основном на практике используют *алгоритм градиентного спуска* [18] – метод, при котором вектор параметров уменьшают на градиент функции потерь, умноженный на гиперпараметр (параметр, не обучающийся в ходе такого обучения) скорости обучения.

Также, чтобы модель не переобучалась (не работала хорошо на данных, на которых она обучается, и плохо – на данных, на которых она не обучается), используют ограничения на функцию потерь – *регуляризацию*. Ограничение, которое используется в данной работе – *L2 регуляризация*. Данное ограничение предполагает, что берется среднее квадратов параметров модели и прибавляется к начальной функции потерь с коэффициентом снижения веса (гиперпараметром).

Детализация используемых методов анализа тональности текстов новостей

В данной работе для *стемминга* мы используем известный *алгоритм Портера* [19]. Векторизация производилась с помощью алгоритма машинного обучения word2vec. Он подробно описывается в работе [20]. Его суть заключается в том, чтобы получать похожие вектора для слов, встречающихся в схожих контекстах. Для каждого слова случайно инициализируются два вектора заданного размера (который представляет собой гиперпараметр) – входной (v) и выходной (v'), из которых после обучения обычно берется среднее, и уже оно используется в векторизации. Далее для каждого слова в тексте берется окно – окрестность размера c из соседних слов. Целью обучения является максимизация среднего логарифма вероятностей по всем словам и комбинациям его со словами в окне. Пусть количество слов в тексте – T , w_t – слово, стоящее на позиции t , v_w – вектор, соответствующий слову w , тогда формула для максимизируемой целевой функции строится следующим образом:

$$\frac{1}{T} \sum_{t=1}^T \sum_{-c \leq j \leq c, j \neq 0} \log p(w_{t+j} | w_t), \quad (11)$$

где $p(w_0 | w_1) = \frac{\exp(v'_w v_{w_1})}{\sum_{w=1}^W \exp(v'_w v_{w_t})}$ – *софтмакс-функция*, часто используемая

в вычислении вероятностей. Однако данный алгоритм непрактичен ввиду низкой скорости вычисления софтмакс-функции. Для упрощения вычисления вероятности на практике используется *метод отрицательной выборки*. Суть этого метода заключается в следующем: сначала вычисляется вероятность,

что слово предскажет слово из окна через *сигмоиду* – $\sigma = \frac{1}{1+e^{-x}}$, а затем из слов, встречающихся в тексте, случайно выбирается k слов (гиперпараметр), не встречающихся в окне, и по ним вычисляется штраф, который прибавляется к логарифму предыдущей вероятности. Таким образом, формула преобразуется следующим образом:

$$\log p(w_0|w_1) = \log \sigma(v'_{w_0} v_{w_1}) + \sum_{i=1}^k E_{w_i \sim P_n(w)} [\log \sigma(-v'_{w_i} v_{w_1})], \quad (12)$$

где $P_n(w)$ – распределение слов в тексте, возведенное в степень $\frac{3}{4}$ и нормализованное.

В качестве *нейронной сети* в данной работе используется комбинация трех типов слоев – *полностью соединенный*, LSTM (Long Short Term Memory, *долгая краткосрочная память*) и CNN (Convolutional Neural Network – *сверточная нейронная сеть*). Выбор нами такой структуры нейронной сети объясняется структурой данных (последовательность данных, для которых нужен LSTM) и популярностью использования CNN для обработки текстовых предложений.

Слой нейронной сети LSTM специально разработан для обработки данных, состоящих из матриц, один из размеров которой неизвестен наперед, т. е., например, матриц векторов слов в тексте. Она подробно описана в работе [21]. Входные данные данного слоя нейронной сети состоят из последовательности векторов одинаковой длины x_t , где t – номер вектора в последовательности. LSTM состоит из одного элемента, который последовательно применяется к последовательным входным данным и двум выходам предыдущего элемента. Всего выходов у этого элемента два – два вектора одинаковой длины, из которых один выбирается в качестве выходных данных на конце. Одним из вариантов является LSTM с *воротами забывания* (LSTM with a forget gate). Элемент данной LSTM можно представить следующими формулами:

$$\begin{aligned} f_t &= \sigma_g(W_f x_t + U_f h_{t-1} + b_f); \quad i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i); \\ o_t &= \sigma_g(W_o x_t + U_o h_{t-1} + b_o); \quad \underline{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c); \quad (13) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \underline{c}_t; \quad h_t = o_t \odot \sigma_h(c_t), \end{aligned}$$

где $x_t \in R^d$ – вектор входных данных для элемента LSTM, $f_t \in (0,1)^h$ – вектор активации ворот забывания (доля долгосрочной памяти предыдущего элемента, которую надо сохранить), $i_t \in (0,1)^h$ – вектор активации ворот входных данных (доля новой долгосрочной памяти, которую необходимо добавить), $o_t \in (0,1)^h$ – вектор активации ворот выходных данных (доля долгосрочной памяти, которую надо перевести в краткосрочную), $h_t \in (-1,1)^h$ – вектор скрытого состояния – один из выходных векторов LSTM, $\underline{c}_t \in (-1,1)^h$ – вектор активации входных данных, $c_t \in R^h$ – вектор состояния (долгосрочная память) – один из выходных векторов элемента LSTM, $W \in R^{h \times d}$, $U \in R^{h \times h}$, $b \in R^h$ – матрицы весов и вектор смещения, которые должны быть обучены на входных

данных. Функции активации: $\sigma_g = \frac{1}{1+e^{-x}}$ – сигмоида, $\sigma_c = \sigma_h = \tanh(x)$ – гиперболический тангенс. h_0 и c_0 инициализируются либо нулями, либо случайными числами, близкими к нулю.

Слой CNN нейронной сети работает в основном на трехмерных тензорах (в том числе и в данной работе). Построение слоя CNN исходит из предположения о том, что методы поиска закономерностей для разных мест во входных данных одинаковы (что, например, имеет место быть в задаче классификации изображений, а также небольших предложений). Его архитектура вводит два типа подслоев: слой *свертки* и слой *пулинга*.

Принцип работы *слоя свертки* следующий: есть трехмерный тензор с осями x , y и канала, он умножается на обучаемый трехмерный тензор поменьше в размерностях x , y (размеры тензора являются гиперпараметрами), но такой же размерностью по каналу в каждой точке, в которой его можно применить, все полученные числа складываются, а также к полученному результату добавляют одно число смещения, которое тоже обучается. Бывает также, что к входным данным добавляют нули по краям (или какие-либо другие числа). Также тензор могут применять не в каждой точке, а через какой-либо шаг по x , y . В основном применяют не один такой тензор, а несколько, иногда даже разных размеров. Таким образом, на выходе может получиться несколько трехмерных тензоров разных размеров.

Слой *пулинга*, в свою очередь, сильно (обычно в два раза по x и в два раза по y) уменьшает размерность входного тензора, однако не меняет размерность канала. Берется некоторый связный подтензор из входных данных из каждого канала в каждой точке, кратной размеру тензора по x , y , соответственно, затем на каждое такое наложение возвращается некоторое число. Различают несколько видов *пулинга*, среди которых: *пулинг по максимальному значению*, в котором оставляют максимальное значение в подтензоре, и *по среднему значению*, в котором возвращают среднее значение в подтензоре. В данной работе использовалась модифицированная нейронная сеть архитектуры CNN из известной статьи [22].

Также в данной работе мы используем модифицированный алгоритм оптимизации вместо обычного градиентного спуска, который обычно никем не применяется. Во-первых, его потребление памяти линейно растет с увеличением количества данных. Во-вторых, он часто застревает в плохих локальных минимумах. В качестве альтернативы ему часто используют SGD – Stochastic Gradient Descent – *стохастический градиентный спуск*, суть которого заключается в том, что на каждом шаге алгоритма считают градиент не всей функции потерь, а только на некоторых случайных данных фиксированного количества – «батчах» (размер такого «батча» – гиперпараметр этого алгоритма).

Также для оптимизации используют и другие алгоритмы, но всегда с применением принципа SGD. Одним из таких алгоритмов оптимизации является Adam. Его суть заключается в том, чтобы вначале считать первые и вторые моменты градиентов: экспоненциально взвешенные суммы градиентов и квадратов градиентов соответственно, а затем вычитать из параметров не градиенты, а моменты, поделенные на корень из максимального второго момента. Такой алгоритм нужен для того, чтобы учитывать некоторые предыдущие градиенты. Если они слишком

большие, то алгоритм считает, что действовать нужно осторожно – обновляться с меньшей скоростью, иначе – пытаться не застрять в плохих (высоких) локальных минимумах (такие обычно с маленьким градиентом в окрестности) с помощью более высокой скорости обучения.

В данной работе используется модификация этого алгоритма – AdamW, для лучшего подсчета градиентов при регуляризации. Дело в том, что классический Adam считает моменты с учетом градиента члена регуляризации. Из-за этого получается, что моменты влияют на степень регуляризации, что приводит к плохой регуляризации. Решение – вынести градиент члена регуляризации за вычисление моментов и не делить его на корень максимума второго момента.

Учет информационного фона в малой DSGE-модели экономики России с адаптивным обучением экономических агентов

DSGE-модель экономики России с малым числом уравнений

Разработанный выше подход учета информационного фона для общего класса линейных (или линеаризованных) моделей с ожиданиями позволяет включать информационный фон в различные варианты структуры неоднородного обучения, учитывая различные модификации PLM-агентов (добавляя в них разные комбинации лагов эндогенных переменных и наблюдаемых шоков) и разные типы неоднородности в обучении, выбирая наиболее подходящую реальному поведению экономических агентов и статистическим данным спецификацию процесса формирования ожиданий применительно к экономике России.

В текущей работе мы рассматриваем применение данного подхода к разработанной нами малой DSGE-модели экономики России с адаптивным обучением [15], которая уже была оценена и использована для проведения сценарных расчетов по экономике России в ИЭОПП СО РАН [23]. В данной модификации модели (с малым числом уравнений) рассматривается упрощенная структура экономики: два типа домашних хозяйств и три вида фирм. В экономике происходит распределение факторов производства: труда, капитала (домашних хозяйств и государства), а также непосредственно самой нефтегазовой продукции.

Структурная неоднородность в экономике задается наличием двух типов домашних хозяйств: «бедных» (нерикардианских), не делающих сбережений на протяжении всей жизни, не имеющих доступа к финансовым рынкам и потребляющих весь свой располагаемый доход каждый период, и дальновидных (рикардианских) домохозяйств, осуществляющих займы друг у друга, занимающих за рубежом, инвестирующих в физический капитал и получающих прибыль. Фирмы в свою очередь делятся на три основных типа, представляющих различные агрегированные секторы: предприятия нефтегазового и обрабатывающего сектора, а также предприятия-ритейлеры. Также предполагается, что все фирмы обрабатывающего сектора принадлежат рикардианским «богатым» домохозяйствам. А предприятия нефтегазовой отрасли владеют не только «богатые» домохозяйства, но и государство и иностранные инвесторы, поэтому между всеми ними распределяется прибыль после выплаты налога в виде дивидендов. Детальное описание модели,

постановку задач агентов и условия равновесия можно найти у Колужнова и Ляхновой [15] и в монографии [23].

Основные уравнения конечной системы логлинеаризованных уравнений, описывающих равновесную динамику этой модели, выглядят следующим образом:

$$R_t = E_t \pi_{t+1} + (\varphi - 1) (\widetilde{C}_t^s - E_t \widetilde{C}_{t+1}^s) = \pi_{t+1}^e + (\varphi - 1) (\widetilde{C}_t^s - \widetilde{C}_{t+1}^{s,e}) - \quad (14)$$

уравнение Эйлера для потребления;

$$\pi_t^{no} = \beta \pi_{t+1}^{no,e} + \frac{(1-\theta)(1-\beta\theta)}{\theta(1+\tau_{no}\beta)} \widetilde{MC}_t^{no} - \quad (15)$$

уравнение новой кейнсианской кривой Филлипс; а

$$\widetilde{MC}_t^{no} = (1 - \alpha_1 - \alpha_2) \widetilde{W}_t + \alpha_1 \widetilde{P}_t^{k,no} - A_t^{no} + \alpha_2 \widetilde{P}_t^{oil} - \quad (16)$$

уравнение предельных издержек;

$$R_t = \rho_R R_{t-1} + (1 - \rho_R) (\varphi_\pi \pi_t + \varphi_{Ne} \widetilde{Ne}_t + \varphi_y \widetilde{Y}_t^{no}) + shock_t^R - \quad (17)$$

правило Тейлора;

$$\widetilde{Ne}_t = \omega_e E_t \widetilde{Ne}_{t+1} + (1 - \omega_e) (\widetilde{Ne}_{t-1} + \omega_i (\pi_t^{ru,target} - \pi_t^{usa,target})) - (R_t - R_t^{usa}) + shock_t^{Ne} - \quad (18)$$

уравнение для номинального обменного курса,

где введены обозначения, представленные в табл. 1. Волны над переменными обозначают, что они выражены в терминах процентного отклонения от устойчивого состояния. Остальные уравнения изложены в статье Колужнова и Ляхновой [15].

Таблица 1

Обозначения переменных в системе уравнений

Table 1

Designations of variables in the system of equations

Обозначение	Смысл
R_t	номинальная внутренняя процентная ставка
R_t^{usa}	номинальная процентная ставка США
π_t	общая инфляция (всех товаров и услуг)
π_t^{no}	инфляция внутреннего конечного товара
$\pi_t^{ru,target}$	долгосрочная цель по инфляции в России
$\pi_t^{usa,target}$	долгосрочная цель по инфляции в США
C_t^s	потребление «богатых» домохозяйств
Y_t^{no}	выпуск не нефтегазовых компаний
W_t	номинальная заработная плата
$P_t^{k,no}$	номинальная цена единицы реального не нефтяного капитала
P_t^{oil}	внутренняя цена на нефть (марки Brent)
Ne_t	номинальный обменный курс (доллара)
A_t^{no}	технологический шок не нефтегазовых фирм
$shock_t^R$	шок монетарной политики
$shock_t^{Ne}$	шок валютного курса

Неоднородное адаптивное обучение в модели с малым числом уравнений

Описанная выше DSGE-модель принадлежит общему классу моделей, рассмотренных выше, а именно представляет собой частный случай вперед смотрящей Модели III. Более точно ее можно представить в виде многомерной линейной структурно однородной модели без лагов эндогенной переменной и с t -датировкой ожиданий:

$$y_t = \alpha + A\widehat{E}_t y_{t+1} + Bw_t \quad (19)$$

и (2) в обозначениях, принятых выше.

Крышка над оператором ожиданий в $\widehat{E}_t y_{t+1}$ показывает, что мы отходим от предположения о рациональных ожиданиях домохозяйств и предполагаем формирование ими ожиданий при помощи адаптивного эконометрического обучения. Более конкретно мы предполагаем, что агенты используют метод RLS как основной алгоритм обучения (т. е. ведут себя как эконометристы, оценивают и пересматривают уравнение регрессии для предсказываемой переменной при помощи метода наименьших квадратов каждый период времени). Уравнения, которые они оценивают, представлены PLM как

$$y_t = a_t + b_t w_t = \varphi_t^T z_t. \quad (20)$$

Векторы y_t , w_t и z_t в соответствии с нашей моделью имеют следующий вид:

$$y_t = \begin{pmatrix} \widetilde{C}_t^s \\ \pi_t \\ \pi_t^{no} \end{pmatrix}; w'_t = (y_{t-1} \quad \widetilde{P}_t^{oil^s} \quad A_t^{no} \quad R_t \quad N_{t-1}); z_t = \begin{pmatrix} 1 \\ w_t \end{pmatrix},$$

$$\text{где } N_t = [N_t^c, N_t^\pi, N_t^{\pi^{no}}]. \quad (21)$$

Более детально, с учетом того, что агенты включают в прогноз переменной лаг только именно этой переменной и только ту метрику новостного фона, которая оказывает влияние именно на нее, PLM, записанные для каждой предсказываемой переменной, выглядят следующим образом:

$$\widetilde{C}_t^s = \alpha_{1,t}^c \times \widetilde{C}_{t-1}^s + \alpha_{2,t}^c \times \widetilde{P}_t^{oil^s} + \alpha_{3,t}^c \times R_t + \alpha_{4,t}^c \times A_t + \alpha_{5,t}^c \times N_{t-1}^c \quad (22)$$

$$\pi_t = \alpha_{1,t}^\pi \times \pi_{t-1} + \alpha_{2,t}^\pi \times \widetilde{P}_t^{oil^s} + \alpha_{3,t}^\pi \times R_t + \alpha_{4,t}^\pi \times A_t + \alpha_{5,t}^\pi \times N_{t-1}^\pi \quad (23)$$

$$\pi_t^{no} = \alpha_{1,t}^{\pi^{no}} \times \pi_{t-1}^{no} + \alpha_{2,t}^{\pi^{no}} \times \widetilde{P}_t^{oil^s} + \alpha_{3,t}^{\pi^{no}} \times R_t + \alpha_{4,t}^{\pi^{no}} \times A_t + \alpha_{5,t}^{\pi^{no}} \times N_{t-1}^{\pi^{no}} \quad (24)$$

Результаты применения подхода учета информационного фона к малой DSGE-модели экономики России с адаптивным обучением

Результаты применения модели анализа тональности

Для получения временных рядов численной метрики, характеризующей информационный фон по отношению к каждой предсказываемой агентами пере-

менной в рассматриваемой модели, нами была построена модель анализа информационного фона на основе новостей. Для написания программ этой модели использовался язык Python 3³.

Поскольку в рассматриваемой модели мы имеем дело с предсказанием квартальных данных, то было необходимо взять новости за как можно больший период. Лучше всего для этой цели подошел источник новостей interfax.ru, потому что у него было доступно больше всего новостей. Из всего набора новостей нами были взяты только новости из раздела «Экономика». По итогу были скачаны новости с 01.01.2010 по 31.12.2021 за каждый день.

Обработка новостей проводилась следующим образом. Вначале из них удалялся стандартный для interfax.ru текст с указанием города, даты и источника. Затем удалялись не нужные нам слова (стоп-слова и английские, не являющиеся названиями компаний) и стоп-символы. После этого производился стемминг алгоритмом Портера [19]. Векторизация текста новостей проводилась алгоритмом `word2vec` с размером вектора 100, обученным на полученных новостях.

Целью применения алгоритма анализа тональности информационного фона стояло получение из всех новостей за квартал единственного числа (от -1 до 1), показывающего влияние новости на конкретный показатель (такой как инфляция, потребление на душу населения). Чтобы модель не переобучалась на таком маленьком количестве данных (данные квартальные), было решено обучить другую нейронную сеть на более частых данных (частотой в один день), которые связаны с инфляцией, и взять полученные за каждый день векторы предпоследнего слоя нейронной сети в качестве входных данных для исходных показателей. В качестве такого дневного показателя был выбран курс американского доллара к рублю ввиду его доступности.

Обучение нейронной сети происходило следующим образом. Вначале текст разделялся на предложения размера в 25 слов (недостающие слова заполнялись векторами из нулей) для того, чтобы можно было их послать на вход CNN. Полученные данные вначале отправлялись в CNN. У CNN была задана следующая архитектура, являющаяся модификацией известной архитектуры для обработки предложений [22], состоящей из трех типов фильтров: фильтр размера $1 \times 4 \times 100$, размера $1 \times 3 \times 100$ и размера $1 \times 2 \times 10$, и таких фильтров по 30 каждого типа. Для каждого фильтра использовалась функция активации ReLU и бралось максимальное значение после фильтра. Затем все значения объединялись, и получался вектор размера 90, который затем отправлялся в полностью соединенный слой размера 90×90 , после чего применялась функция активации ReLU.

На следующем шаге полученные векторы предложений объединялись в различные матрицы текстов и отправлялись в LSTM архитектуры по умолчанию с размерностью выходного вектора 90. Затем полученные векторы текстов отправлялись в два полностью соединенных слоя 90×90 с ReLU после каждого. После

³ Для получения HTML из новостей использовалась библиотека BeautifulSoup4. Для алгоритма Портера использовалась библиотека NLTK. Для алгоритма `word2vec` использовалась библиотека `gensim`. Для работы с табличными данными и очистки текста использовалась библиотека `pandas`. Для построения графиков использовалась библиотека `matplotlib`. Для построения нейронных сетей использовалась библиотека `PyTorch`. Для подбора гиперпараметров использовался алгоритм ASHA из библиотеки `ray[tune]`.

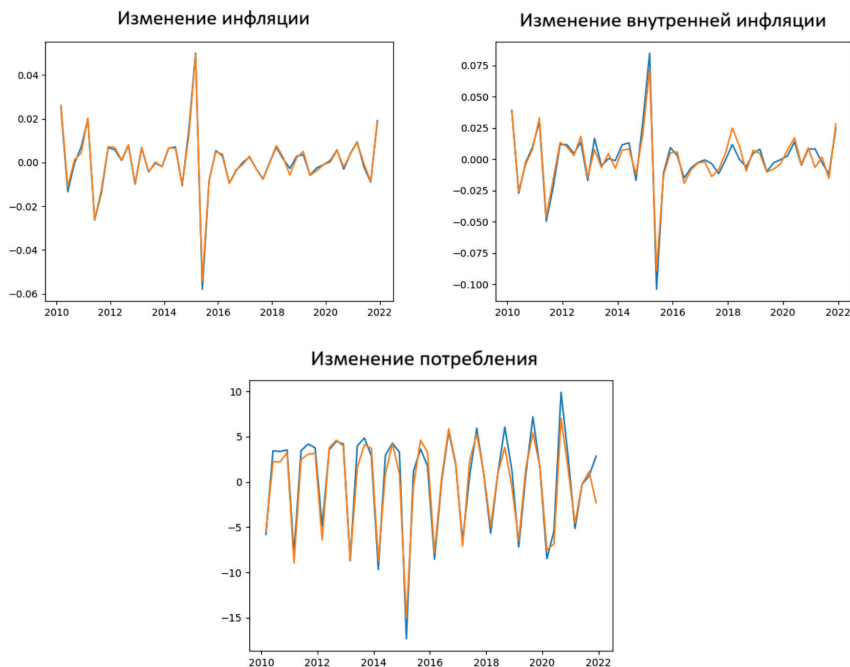


Рис. 1. Прогноз изменения инфляции, внутренней инфляции и потребления при помощи информационного фона

Fig. 1. Forecasting changes in inflation, domestic inflation and consumption using the information background

этого полученные векторы текстов объединялись в различные матрицы по дням и отправлялись в LSTM архитектуры по умолчанию с размерностью выходного вектора 90.

На следующем этапе полученные векторы текстов отправлялись в два полностью соединенных слоя – первый размером 90×90 с ReLU после него. Второй – размером 90×1 без функции активации – для регрессии. Обучалась такая нейронная сеть с помощью AdamW со скоростью обучения 10^{-4} и снижением веса 10^{-1} . Для дальнейшего обучения брался вектор размера 90 для дня с предпоследнего слоя.

Для предсказания всех целевых показателей использовалась вторая нейронная сеть, с одинаковой архитектурой модели для предсказания всех переменных модели. Для получения метрик новостного фона от -1 до 1 с сохранением знака, значения предсказываемых переменных делились на $\max\{|\min(\text{target})|, |\max(\text{target})|\}$, где target – фактические значения предсказываемых переменных. Для каждого квартала все векторы из предыдущего пункта объединялись по порядку в матрицу для квартала. Затем эта матрица отправлялась в LSTM стандартной архитектуры с размером выходного вектора 90. После этого вектор отправлялся в полностью соединенный слой размера 90×1 , а затем – в гиперболический тангенс – для предсказания значений от -1 до 1 . Обучалась такая нейронная сеть с помощью AdamW со скоростью обучения 10^{-5} и снижением веса 10^{-1} .

То, что полученные нами в результате применения данного метода метрики для информационного фона по каждой переменной хорошо описывают изменение каждой переменной, видно на представленных графиках (рис. 1). По оси абсцисс отложены кварталы, за которые собирались новости, по оси ординат – показатели изменения каждой предсказываемой переменной в исходных единицах измерения. Синей линией показаны фактические данные, желтым – предсказания, полученные на основе применения модели анализа информационного фона.

Результаты учета информационного фона в адаптивном обучении агентов, ретроспективные прогнозы

Полученные метрики информационного фона были использованы нами в моделях формирования и пересмотра ожиданий (адаптивного обучения) соответствующих переменных экономическими агентами в малой DSGE-модели экономики России с адаптивным обучением.

Для проверки гипотезы об улучшении прогнозной точности за счет учета новостного фона нами были построены ретроспективные прогнозы для ключевых переменных для двух случаев – с учетом и без учета новостей. Мы смотрим, какие оценки оказались ближе к реальной экономической динамике при разных модификациях модели (с точки зрения учета новостей при формировании агентами ожиданий).

Статистика по всем используемым в системе логлинеаризованных уравнений переменным была взята с первого квартала 2010 г. по четвертый квартал 2022 г., а в качестве временного интервала для тестирования был взят период с первого квартала 2015 г. по четвертый квартал 2022 г. Полученные по основным макроэкономическим показателям результаты отражены на шести графиках рис. 2. Потребление выражено в млрд руб; инфляция – в %, QoQ⁴; выпуск (ВВП) – в ценах 2011 г., млрд руб.; рабочая сила – в млн человек.

Видно, что наибольшее влияние учет новостного фона оказал на такие показатели, как процентная ставка, курс доллара и инфляция. В подтверждение этого визуального результата были рассчитаны такие метрики качества, как MAE и MSE для каждой модификации модели (табл. 2).

Таблица 2

Оценка качества прогноза ключевых макропеременных

Table 2

Assessment of the quality of forecasting of key macro variables

MSE	C	R	infl	Ne	Y	L
без новостей	423969	5,0	4,7	84,2	4637	5,8
с новостями	412159	4,6	4,4	63,7	4653	5,7
MAE	C	R	infl	Ne	Y	L
без новостей	567	1,6	1,5	6,8	97	1,9
с новостями	555	1,3	1,3	4,6	90	1,8

⁴ Quarter-over-Quarter, т. е. квартал к предыдущему кварталу.

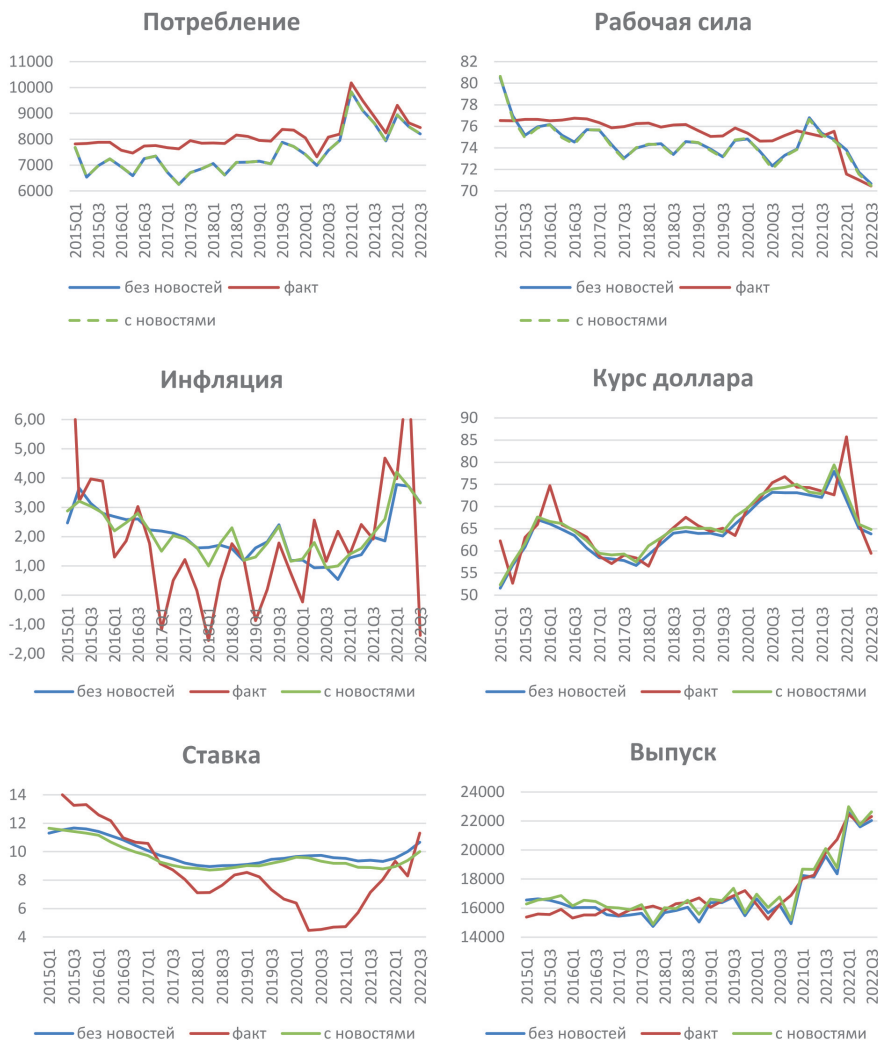


Рис. 2. Фактические и прогнозные траектории для основных макроэкономических показателей в 2015–2022 гг.

Fig. 2. Actual and forecast trajectories for key macroeconomic indicators in 2015–2022

Для всех показателей значения MAE и MSE оказались меньше для модели, где при построении ожиданий учитываются новости, нежели чем без их учета.

Прогнозы ключевых макроэкономических показателей на 2023 г.

На предыдущем шаге было показано, что построенная нами DSGE-модель с адаптивным обучением, учитывающая новостной фон, позволяет получать более

точные расчеты (прогнозы). Поэтому для дальнейшего прогнозирования основных макроэкономических показателей, например, на один год вперед (на 2023 г.), будем использовать именно ее.

Основное предположение, которое было сделано при построении базового прогноза, заключается в том, что большинство наблюдаемых переменных сохранят схожую траекторию движения (экстраполяция данных с учетом сезонности), в то же время несколько уменьшится инфляция импортных товаров и мировая цена нефти, а новостной фон будет нейтральным.

В качестве основных макроэкономических показателей, для которых были получены прогнозы, брались те, которые обычно прогнозирует Центральный банк – выпуск, потребление, инфляция, процентная ставка и курс USD.

Таблица 3

Прогноз макроэкономических показателей на 2023 г.

Table 3

Forecast of macroeconomic indicators for 2023

Показатель	Прогноз на 2023 г.
Изменение инфляции за год	7,2 %
Средняя за год процентная (ключевая) ставка	8,2 %
Изменение ВВП за год	3 %
ВВП, IV квартал к IV кварталу прошлого года	1,1 %
Изменение расходов домашних хозяйств за год	5 %
Курс доллара на конец года	100,8 рублей за доллар
Среднегодовой курс доллара	85 рублей за доллар

Согласно результатам, отраженным в табл. 3, по итогу 2023 г. инфляция составит около 7,3 %, среднегодовая процентная ставка будет примерно 8,2 %, а среднегодовой курс доллара составит 85 рублей за доллар. При этом ВВП вырастет за год на 3 % и расходы домохозяйств – на 5 %. Эти полученные нами прогнозы согласуются с теми среднесрочными прогнозами по базовому сценарию, которые публикует Банк России⁵.

Сценарное прогнозирование наиболее зависимых от новостей показателей на 2023 г.

Помимо базового сценария, когда новости нейтральны, рассмотрим еще два сценария: при исключительно положительных новостях (метрика новостного фона в этом случае равна 1) и исключительно отрицательных (метрика новостного

⁵ Например, https://cbr.ru/Collection/Collection/File/46328/forecast_230915.pdf

фона равна -1). Результаты по наиболее зависимым от новостного фона переменным (курсу доллара, инфляции и ставке) отражены в табл. 4.

Таблица 4

Результаты сценарных прогнозов на 2023 год

Table 4

Results of scenario projections for 2023

Сценарий	Курс доллара	Инфляция, %	Ставка, %
При положительных новостях	84 руб. за долл.	5,1	7,3
При нейтральных новостях	85 руб. за долл.	7,2	8,2
При негативных новостях	86 руб. за долл.	9,3	9,1

В случае положительного новостного фона инфляция будет существенно меньше – около 5,1 %, аналогична ситуация по ставке – она составит 7,3 %, курс доллара же снизится незначительно – будет примерно 84 рубля за доллар. При отрицательном новостном фоне ситуация противоположная: незначительное увеличение курса до 86 рублей за доллар и более существенный рост инфляции (до 9,3 %) и ставки (до 9,1 %).

Заключение

Итогом данной работы является разработка универсального метода учета информационного фона для широкого класса моделей с адаптивным обучением экономических агентов (включая DSGE-модели) и построение с использованием методов машинного обучения модели анализа информационного фона на основе потока новостей для предсказания эндогенных переменных модели, переводящей поток новостей в характеризующую этот фон численную метрику. Работа метода учета информационного фона и модель анализа информационного фона были протестированы на малой DSGE-модели экономики России с адаптивным обучением, разработанной в ИЭОПП СО РАН [23].

На примере этой модели было показано, что включение информационного фона в модель формирования ожиданий адаптивно обучающихся агентов улучшает соответствие данных, имитируемых моделью, экономической статистике, что позволяет использовать эту модель для прогнозирования макроэкономических показателей. Включение информационного фона позволяет рассматривать различные сценарии развития экономики при разной окраске будущего информационного фона: позитивной, негативной, нейтральной. Полученные в работе результаты показывают, что управление новостным потоком оказывает влияние на функционирование экономики и может потенциально использоваться для оценки эффектов влияния информационного фона (в том числе специально создаваемого и поддерживаемого регулятором и экономическими властями) на макроэкономическую динамику.

Универсальность метода, предложенного в работе, позволяет распространить его применение на широкий ряд DSGE-моделей, используемых центральными банками большинства стран мира, в том числе Банком России.

Список литературы

1. **Christoffel K., Coenen G., and Warne A.** The New Area-Wide Model of the Euro Area: A Micro-Founded Open-Economy Model for Forecasting and Policy Analysis // ECB Working Paper Series. 2008. No. 944. 124 p. URL: <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp944.pdf>
2. **Christiano L., Rostagno M., and Motto R.** Financial factors in economic fluctuations // ECB Working Paper Series. 2010. No. 1192.
3. **Edge R., Kiley M., and Laforge J.-P.** A Comparison of Forecast Performance Between Federal Reserve Staff Forecasts, Simple Reduced-Form Models, and a DSGE Model // Journal of Applied Econometrics. 2010. Vol. 25. P. 720–754.
4. **Chung H. T., Kiley M. T., and Laforge J.-P.** Documentation of the Estimated, Dynamic, Optimization-based (EDO) Model of the U.S. Economy: 2010 Version // Finance and Economics Discussion Series Working Paper. 2010. No. 19.
5. **Fenton P., Murchison S. C.** ToTEM: The Bank of Canada's New Projection and Policy-Analysis Model // Bank of Canada Review. 2006. Vol. 2006. P. 5–18.
6. **Dorich J., Johnston M., Mendes R., Murchison S., and Zhang Y.** ToTEM II: An Updated Version of the Bank of Canada's Quarterly Projection Model // Canadian Economic Analysis Department. Technical Report 100. Bank of Canada. 2013. No. 100.
7. **Harrison R., Nikolov K., Quinn M., Ramsay G., Scott A. Thomas R.** The Bank of England Quarterly Model // Bank of England Publications, 2005.
8. **Burgess S., Fernandez-Corugedo E., Groth C., Harrison R., Monti F., Theodoridis K., and Waldron M.** The Bank of England's Forecasting Platform: COMPASS, MAPS, EASE and the Suite of Models // Bank of England working Paper. 2013. No. 471.
9. **Brubakk L., Anders T., Maih J., Olsen K., Ostnor M.** Finding NEMO: Documentation of the Norwegian economy model // Norwegian Central Bank, Staff Memo. 2006. No. 2006-6. 85 p. URL: <https://www.econstor.eu/bitstream/10419/210178/1/nb-staff-memo2006-06.pdf>
10. **Pesenti P.** The Global Economy Model: Theoretical Framework // IMF Staff Papers. 2008. Vol. 55. No. 2. P. 243–284.
11. **Крепцев Д. А., Селезнев С. М.** DSGE-модель российской экономики с банковским сектором // Серия докладов об экономических исследованиях ЦБ РФ. 2017. № 27. С. 1–82.
12. **Крепцев Д. А., Селезнев С. М.** DSGE-модели российской экономики с малым количеством уравнений // Серия докладов об экономических исследованиях ЦБ РФ. 2016. № 12. С. 1–53.
13. **Sargent T. J.** Bounded Rationality in Macroeconomics. Oxford; N.Y.: Oxford University Press, Clarendon Press, 1993.

14. **Evans G. W., Honkapohja S.** Learning and Expectations in Macroeconomics // Princeton, NJ.: Princeton University Press. 2001.
15. **Колюжнов Д. В., Ляхнова М. В.** Малая DSGE-модель экономики России с неоднородным адаптивным обучением // Мир экономики и управления. – 2022. Т. 22 (3). С. 66–87.
16. **Богомолова А. С., Колюжнов Д. В.** Экономическая динамика при неоднородном адаптивном обучении: условия стабильности в терминах агрегированной экономики // Мир экономики и управления. 2020. Т. 20, № 1. С. 128–153.
17. **Giannitsarou Ch.** Heterogeneous learning // Review of Economic Dynamics. 2003. Vol. 6. pp. 885–906.
18. **Goodfellow I., Bengio Y., Courville A.** Deep Learning // MIT Press, 2016.
19. **Porter M. F.** An algorithm for suffix stripping // Program: Electronic Library and Information Systems. 1980. Vol. 14 (3), pp. 130–137.
20. **Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J.** Distributed representations of words and phrases and their compositionality // Proceedings of the 26th International Conference on Neural Information Processing Systems. 2013. pp. 3111–3119.
21. **Hochreiter, S., Schmidhuber, J.** Long Short-Term Memory // Neural computation. 1997. Vol 9 (8), pp. 1735–1780.
22. **Kim Y.** Convolutional Neural Networks for Sentence Classification // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2014. pp. 1746–1751.
23. **Колюжнов Д. В., Ляхнова М. В.** DSGE-модели для краткосрочного прогнозирования экономики России // Модели и методы прогнозирования: Азиатская Россия в экономике страны / Под ред. А. О. Баранова, В. И. Суслова; Институт экономики и организации промышленного производства Сибирского отделения РАН. Новосибирск : Изд-во ИЭОПП СО РАН, 2023. Гл. 2.1. С. 16–87.

References

1. **Christoffel K., Coenen G., and Warne A.** The New Area-Wide Model of the Euro Area: A Micro-Founded Open-Economy Model for Forecasting and Policy Analysis // ECB Working Paper Series. 2008. No. 944. 124 p. URL: <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp944.pdf>
2. **Christiano L., Rostagno M., and Motto R.** Financial factors in economic fluctuations // ECB Working Paper Series. 2010. No. 1192.
3. **Edge R., Kiley M., and Laforte J.-P.** A Comparison of Forecast Performance Between Federal Reserve Staff Forecasts, Simple Reduced-Form Models, and a DSGE Model // Journal of Applied Econometrics. 2010. Vol. 25. P. 720–754.
4. **Chung H. T., Kiley M. T., and Laforte J.-P.** Documentation of the Estimated, Dynamic, Optimization-based (EDO) Model of the U.S. Economy: 2010 Version // Finance and Economics Discussion Series Working Paper. 2010. No. 19.
5. **Fenton P., Murchison S. C.** ToTEM: The Bank of Canada's New Projection and Policy-Analysis Model // Bank of Canada Review. 2006. Vol. 2006. P. 5–18

6. **Dorich J., Johnston M., Mendes R., Murchison S., and Zhang Y.** ToTEM II: An Updated Version of the Bank of Canada's Quarterly Projection Model // Canadian Economic Analysis Department. Technical Report 100. Bank of Canada. 2013. No. 100.
7. **Harrison R., Nikolov K., Quinn M., Ramsay G., Scott A. Thomas R.** The Bank of England Quarterly Model // Bank of England Publications, 2005.
8. **Burgess S., Fernandez-Corugedo E., Groth C., Harrison R., Monti F., Theodoridis K., and Waldron M.** The Bank of England's Forecasting Platform: COMPASS, MAPS, EASE and the Suite of Models // Bank of England working Paper. 2013. No. 471.
9. **Brubakk L., Anders T., Maih J., Olsen K., Ostnor M.** Finding NEMO: Documentation of the Norwegian economy model // Norwegian Central Bank, Staff Memo. 2006. No. 2006-6. 85 p. URL: <https://www.econstor.eu/bitstream/10419/210178/1/nb-staff-memo2006-06.pdf>
10. **Pesenti P.** The Global Economy Model: Theoretical Framework // IMF Staff Papers. 2008. Vol. 55. No. 2. P. 243–284.
11. **Krepcev D., Seleznev S.** DSGE-model' rossijskoj ekonomiki s bankovskim sektorom // Seriya dokladov ob ekonomicheskikh issledovaniyakh CB RF. 2017. No. 27. pp. 1–82.
12. **Krepcev D., Seleznev S.** DSGE-model' rossijskoj ekonomiki s malym kolichestvom uravnenij // Seriya dokladov ob ekonomicheskikh issledovaniyakh CB RF. 2016. No. 12. pp. 1–53.
13. **Sargent T. J.** Bounded Rationality in Macroeconomics. Oxford; N.Y.: Oxford University Press, Clarendon Press, 1993.
14. **Evans G. W., Honkapohja S.** Learning and Expectations in Macroeconomics // Princeton, NJ.: Princeton University Press. 2001.
15. **Kolyuzhnov D. V., Lyahnova M. V.** Small DSGE Model of the Russian Economy with Heterogeneous Adaptive Learning. World of Economics and Management, 2022, vol. 22, No. 3, pp. 66–87. (in Russ.)
16. **Bogomolova A. S., Kolyuzhnov D. V.** Economic Dynamics under Heterogeneous Adaptive Learning: Aggregate Economy Sufficient Conditions for Stability World of Economics and management, 2020, vol. 20, no. 1, pp. 128–153. (in Russ.)
17. **Giannitsarou Ch.** Heterogeneous learning // Review of Economic Dynamics. 2003. Vol. 6. pp. 885–906.
18. **Goodfellow I., Bengio Y., Courville A.** Deep Learning // MIT Press, 2016.
19. **Porter M. F.** An algorithm for suffix stripping // Program: Electronic Library and Information Systems. 1980. Vol. 14 (3), pp. 130–137.
20. **Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J.** Distributed representations of words and phrases and their compositionality // Proceedings of the 26th International Conference on Neural Information Processing Systems. 2013. pp. 3111–3119.
21. **Hochreiter, S., Schmidhuber, J.** Long Short-Term Memory // Neural computation. 1997. Vol 9 (8), pp. 1735–1780.

22. **Kim Y.** Convolutional Neural Networks for Sentence Classification // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2014. pp. 1746–1751.
23. **Kolyuzhnov D. V., Lyahnova M. V.** DSGE-models for short-term forecasting of Russian economy // Forecasting Models and Methods: Asian Russia in the Country's Economy / edited by A.O. Baranov, V.I. Suslov. Novosibirsk: SB RAS: IEIE SBRAS. 2023. Ch. 2.1. pp. 16–87. (in Russ.)

Информация об авторах

Дмитрий Васильевич Колюжнов, PhD, доцент экономического факультета Новосибирского национального исследовательского государственного университета; научный сотрудник Института экономики и организации промышленного производства СО РАН
Scopus ID 55940049500

Егор Дмитриевич Колюжнов, студент 4-го курса бакалавриата факультета информационных технологий Новосибирского национального исследовательского государственного университета

Маргарита Валерьевна Ляхнова, аспирант Института экономики и организации промышленного производства СО РАН; главный экономист Банка России

Information about the Authors

Dmitriy V. Kolyuzhnov, PhD (Economics), Associate Professor, Novosibirsk State University; Researcher, Institute of Economics and Industrial Engineering SB RAS
Scopus ID 55940049500

Egor D. Kolyuzhnov, 4th year Undergraduate Student, Information Technologies Department, Novosibirsk State University

Margarita V. Lyakhnova, Postgraduate Student, Institute of Economics and Industrial Engineering SB RAS; Chief economist, Bank of Russia

*Статья поступила в редакцию 02.11.2023;
одобрена после рецензирования 25.11.2023; принята к публикации 25.11.2023*

*The article was submitted 02.11.2023;
approved after reviewing 25.11.2023; accepted for publication 25.11.2023*