

Возможности использования высокочастотных данных по онлайн ценам для прогнозирования официальной инфляции

Современные возможности сбора и обработки информации позволяют использовать массивы высокочастотных данных для решения различных исследовательских задач. Примером такого массива данных являются цены онлайн-ритейлеров, доступные с ежедневной частотой по широкому перечню товаров.

Согласно актуальным исследованиям, учет данных цен онлайн-ритейлеров может быть полезен для улучшения прогноза официальной инфляции (Macias P., Stelmasiak D., Szafranek K. 2022), (Hull I. et al 2017). Онлайн-данные характеризуются более высокой частотой измерения, а значит содержат больше информации (Cavallo и Rigobon 2016). Данные доступны в режиме реального времени, и уже «сегодня» содержат информацию, которая будет отражена в данных официальной статистики с лагом (Aparicio D., Bertolotto M. I. 2020). Также для онлайн-цен характерна меньшая жесткость (Божечкова и Евсеев 2020), и постепенное изменение цен формирует ценовую тенденцию раньше, чем официальный ИПЦ.

Классические модели прогноза инфляции либо не учитывают данные об онлайн ценах, либо учитывают их в агрегированном виде (например, модели векторной авторегрессии единой частоты данных). Учет высокочастотной информации о ценах при прогнозировании официальной инфляции возможен в моделях смешанной частоты данных, в частности, в MIDAS-моделях и MF-VAR моделях. Краткий обзор преимуществ и недостатков этих методов представлен в таблице 1.

Автором проводится проверка гипотезы, что учет информации об онлайн ценах (как в агрегированном, так и в высокочастотном виде) способен улучшить прогноз официальной инфляции. В качестве прогнозируемой переменной использованы месячные данные официального ИПЦ на продовольственные товары в Москве в период с февраля 2019 г. по сентябрь 2022 гг. (44 наблюдения). В качестве переменной интереса использован индекс цен на продовольственные товары московских онлайн-ритейлеров, рассчитываемый сотрудниками ИПЭИ РАНХиГС с февраля 2019 г. по настоящее время (ноябрь 2022 г.) с ежедневной частотой (1360 наблюдений на момент расчетов). Динамика этих показателей в течение рассматриваемого периода представлена на рисунке 1.

Таблица 1 – Эконометрические методы прогноза офлайн-инфляции с использованием данных по онлайн-ценам (резюме)

Метод	Работа	Преимущества метода	Недостатки метода
ARMA-модель	(Vicente и Pereira 2022)	Результаты моделей используются для сравнения с результатами, полученными другими методами	Модель не содержит данные об онлайн ценах
VAR-модель	(Aparicio D., Bertolotto M. I. 2020)	Позволяет учесть содержательные факторы зависимой переменной и агрегированные данные онлайн ИПЦ. Позволяет учесть проблему эндогенности	В результате агрегирования высокочастотного регрессора до переменной самой низкой частоты происходит потеря информации
MIDAS-модель	(Ghysels, Santa-Clara и Valkanov 2004)	Позволяет использовать высокочастотные данные для предсказания менее частотных данных без потери информации в результате агрегирования	В модели без ограничений высок риск мультиколлинеарности. Результат прогноза чувствителен к выбору функции ограничения. Проблема эндогенности
MF-VAR-модель	(Schorfheide и Song 2013)	Позволяет учесть содержательные факторы зависимой переменной и высокочастотные данные онлайн ИПЦ. Позволяет учесть проблему эндогенности	Учет содержательных факторов может потребовать снижения частоты высокочастотной переменной

Примечание – Источник: составлено автором.

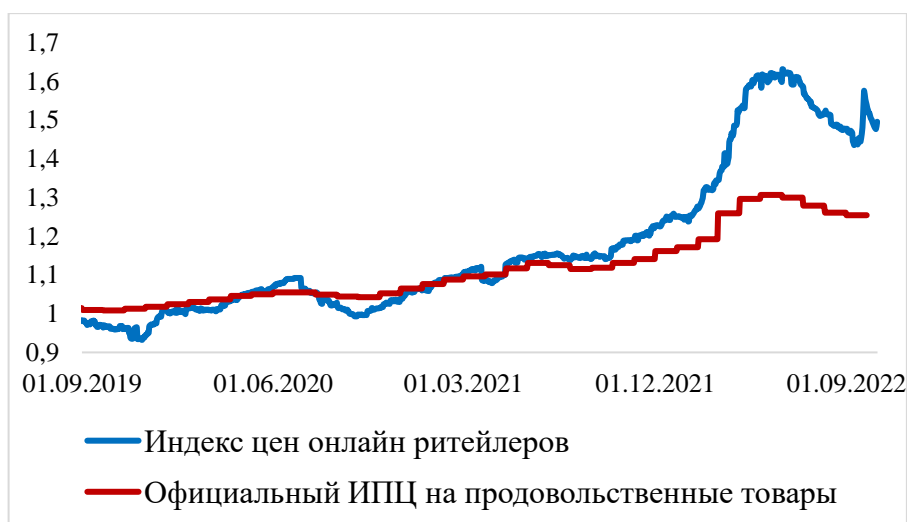


Рисунок 1 – Исследуемый массив данных онлайн индекса цен (ежедневные данные) и официального ИПЦ (месячные данные) в 2019-2022 гг.

Источник: данные сотрудников ИПЭИ РАНХиГС и данные Росстата.

С опорой на обзор методов (таблица 1), автором построены следующие модели официального ИПЦ на продовольственные товары:

А) ARMA-модель (используется в качестве технического бенчмарка):

$$p_t = \gamma + \sum_{i=1}^p \alpha_i p_{t-i} + \sum_{i=1}^q \beta_i \varepsilon_{t-i} + \varepsilon_t, \quad (1)$$

где p_t – это официальная инфляция в период t ;

p – количество лагов AR-процесса;

q – количество лагов MA-процесса;

Б) Стандартная SVAR-X-модель, учитывающая классические факторы инфляции (курс рубля к доллару США, ставка процента MIACR; в качестве экзогенной переменной использован индекс цен ФАО на продовольственные товары).

В) Стандартная SVAR-X-модель, учитывающая информацию об онлайн индексе цен:

$$B_0 Y_t = A_1 X_{t-1} + \dots + A_p X_{t-p} + B_1 Y_{t-1} + \dots + B_p Y_{t-p} + u_t, \quad (2)$$

$$\varepsilon_t = B_0^{-1} u_t, \quad (3)$$

где X_t – вектор экзогенных макроэкономических переменных;

Y_t – вектор эндогенных макроэкономических переменных;

p – количество лагов модели.

Модели (А) и (Б) не содержат информацию об онлайн ценах и построены для оценки характеристик классических прогнозных моделей; модель (В) содержит информацию об онлайн ценах в агрегированном виде (с месячной частотой);

Г) MIDAS-модели с четырьмя стандартными функциями взвешивания (PDL-функция, ступенчатая, Beta-функция и экспоненциальная функция Алмона). Модели (Г) содержат информацию об онлайн ценах в исходном виде (с ежедневной частотой):

$$A(L)\pi_t = c + B(L^{1/m})x_t^{(m)} + \varepsilon_t, \quad (4)$$

где π_t – официальная инфляция, измеряемая с месячной частотой $\{\pi_t, t \in Z\}$,

x_t – онлайн-инфляция, измеряемая с m раз большей частотой $\{x_t^{(m)}, t \in Z\}$.

Д) MF-VAR-модель. Модель (Д) содержит информацию об онлайн ценах в частично агрегированном виде (с недельной частотой):

$$y_t = Ax_t + \varepsilon_t, \quad (5)$$

где y_t, x_t – векторы переменных модели, в том числе переменные, измеряемые с месячной и недельной частотой.

В результате построений получены вневыборочные прогнозы официального ИПЦ на продовольственные товары в Москве на 1 месяц и на 3 месяца вперед. Сравнение прогнозов моделей произведено с опорой на показатель среднеквадратичной ошибки RMSE. Резюме характеристик полученных прогнозов представлено на рисунке 2.

Дата прогноза	Прогнозы на 1 месяц вперед				Дата прогноза	Прогнозы на 3 месяца вперед			
	Группа MIDAS	Группа VAR (online)	VAR	SARIMA		Группа MIDAS	Группа VAR (online)	VAR	SARIMA
окт.21	0,0095	0,0032	0,0031	0,01	окт.21	0,0162	0,0206	0,0231	0,0082
ноя.21	0,0088	0,001	0,0013	0,01	ноя.21	0,0074	0,0061	0,0045	0,0129
дек.21	0,0015	0,0039	0,0028	0,01	дек.21	0,0109	0,0074	0,0041	0,0173
январь.22	0,0087	0,0084	0,0096	0,01	январь.22	0,0104	0,007	0,0073	0,01
фев.22	0,0061	0,0038	0,007	0,0063	фев.22	0,0062	0,0077	0,0079	0,007
мар.22	0,0093	0,015	0,017	0,01	мар.22	0,0103	0,0116	0,0146	0,0058
апр.22	0,0402	0,0163	0,0648	0,0559	апр.22	0,0222	0,0332	0,0353	0,0452
май.22	0,042	0,0484	0,0276	0,04	май.22	0,0306	0,0237	0,047	0,0727
июнь.22	0,0223	0,0211	0,0139	0,0248	июнь.22	0,0297	0,0629	0,0639	0,0757
июль.22	0,0276	0,0167	0,0171	0,0178	июль.22	0,0571	0,0371	0,0499	0,0264
авг.22	0,0109	0,0096	0,0087	0,0144	авг.22	0,0238	0,0503	0,043	0,0508
сентябрь.22	0,0035	0,0071	0,0024	0,011	сентябрь.22	0,0218	0,0331	0,0371	0,0318
октябрь.22	0,0027	0,0034	0,0028	0,0111	октябрь.22	0,01	0,0098	0,0049	0,0117

Рисунок 2 – Значение RMSE в группах прогнозных моделей официального ИПЦ на продовольственные товары на 1 и 3 месяца вперед

Примечание – в качестве даты прогноза указана дата, на которую сделан прогноз. Красным цветом обозначены прогнозы с наибольшей ошибкой, синим – с наименьшей ошибкой. Источник: рассчитано автором.

Согласно полученным результатам, SARIMA-модели характеризуются наибольшей ошибкой прогноза. VAR-модели, содержащие информацию об онлайн ценах, в среднем характеризуются меньшими значениями RMSE, чем VAR-модель, не использующая эту информацию. Модели MIDAS учитывают изменение динамики цен на ежедневной основе, что может быть востребовано для прогноза в кризисных условиях. Модели MF-VAR учитывают эндогенность макроэкономических переменных, в некотором объеме учитывают информацию высокочастотных переменных и имеют лучшую содержательную интерпретируемость. Таким образом, автор приходит к выводу, что включение высокочастотных данных в модель прогноза официальной инфляции способно улучшить качество краткосрочного прогноза, и такие прогнозные техники могут быть полезны для оперативных целей экономической политики в России, в том числе и в кризисных условиях.