

Санкт–Петербургский государственный университет

Дорофеев Олег Артёмович

Выпускная квалификационная работа

**Прогнозирование временных рядов с множеством факторов в
макроэкономике и проверка гипотез о зависимости между факторами**

Уровень образования: бакалавриат

Направление 02.03.02 «Фундаментальная информатика и
информационные технологии»

Основная образовательная программа СВ.5003.2020

«Программирование и информационные технологии»

Научный руководитель:

старший преподаватель кафедры
диагностики функциональных систем,
кандидат физико-математических наук

Парфенов Андрей Павлович

Рецензент:

доктор технических наук,
профессор кафедры математической теории игр и
статистических решений

Буре Владимир Мансурович

Санкт-Петербург

2024

Saint-Petersburg State University

Dorofeev Oleg Artemovich

Final qualifying work

**Multivariate macroeconomic time series forecasting and testing hypotheses
about the dependence between factors**

Education level: bachelor's degree

Education program 02.03.02 «Fundamental computer science and
information technology»

The main educational program CB.5003.2020

"Programming and Information Technology"

Scientific supervisor:

Senior Lecturer at the Department
of Functional Systems Diagnostics,

Candidate of Physical and Mathematical Sciences

Parfenov A. P.

Reviewer:

Doctor of Technical Sciences,

Professor of the Department of Mathematical Game Theory and

Statistical Solutions

Bure V. M.

Saint-Petersburg

2024

Оглавление

Введение.....	4
Актуальность работы.....	6
Цели работы.....	7
Глава 1. Данные и теория.....	8
1.1 Описание показателей.....	8
1.2 Используемый инструментарий.....	10
1.3 Первичный анализ и обработка данных.....	11
1.4 Стационарность.....	14
1.5 Автокорреляция.....	17
1.6 Матрицы корреляций.....	18
1.7 Поправка Бонферрони.....	20
1.8 Тест Гренджера.....	21
Глава 2. Прогнозирование и исследование.....	23
2.1 Модели прогнозирования.....	23
2.2 Ошибки.....	24
2.3 Тестирование моделей.....	25
2.4 Прогнозирование на посткризисный период.....	27
2.5 Нормальное распределение.....	27
2.7 U-критерий Манна-Уитни.....	28
Прогнозирование на будущее.....	32
Заключение.....	34
Список использованной литературы.....	36
Приложение.....	37

Введение

Макроэкономические показатели на протяжении многих лет постоянно отслеживаются правительственными учреждениями разных стран с целью обеспечения надлежащего функционирования экономики. Экономика развивается с течением времени и подвержена периодическим, а иногда и крупным непредвиденным потрясениям. Кроме того, между различными факторами часто существуют тесные взаимосвязи, которые необходимо учитывать при прогнозировании. К примеру, рост уровня безработицы вызывает сокращение потребительских расходов, а изменения в денежно-кредитной политике влияют на инфляцию и инвестиции.

Последним крупным кризисом, действительно повлиявшим на весь мир и вызвавшим глобальные экономические проблемы, стала пандемия Covid-19. Развернувшийся кризис оказал беспрецедентное воздействие на мировую экономику, вызвав резкий спад в различных секторах, от туризма и розничной торговли до производства и энергетики. Вследствие введения строгих карантинных мер, закрытия границ и ограничений на передвижение, многие компании столкнулись с серьёзными финансовыми трудностями, что привело к массовым увольнениям.

Экономический шок, вызванный пандемией, проявился в различных формах: от сокращения ВВП и роста безработицы до изменения потребительских привычек и нарушений в глобальных цепочках поставок. Эти изменения привели к необходимости адаптации экономической политики и мер поддержки со стороны правительств и центральных банков. Для понимания глубины и продолжительности этих изменений важно рассмотреть структурные сдвиги, которые могли произойти в экономике после 2020 года.

В данном исследовании поднимается вопрос о наличии структурных изменений в экономике после кризисного шока. Для этого мы задействуем различные макроэкономические показатели США и рассмотрим их в двух временных периодах:

- I. До января 2020 года – объявления Всемирной организацией здравоохранения о чрезвычайной ситуации международного значения;
- II. С января 2021 года – снятия большинства карантинных ограничений и нормализации ВВП США.

Анализ этих двух периодов позволит выявить структурные изменения, произошедшие в экономике США в результате кризиса, и понять, как экономика адаптировалась к новым условиям. Это, в свою очередь, поможет разработать рекомендации по дальнейшему развитию экономической политики и мер поддержки, направленных на обеспечение устойчивого экономического роста и предотвращение будущих кризисов.

Актуальность работы

Актуальность данного исследования обусловлена необходимостью глубокого анализа структурных изменений в экономике на примере США после кризисного шока, вызванного пандемией. На основе анализа экономики США можно предположить, какие структурные изменения происходят и в экономиках других стран. Понимание этих изменений позволит разработать более эффективные меры для предотвращения подобных кризисов в будущем и обеспечения устойчивого экономического роста. Несмотря на обилие исследований, посвященных экономическим последствиям пандемии, мало кто использовал методологию прогнозирования для оценки различий между периодами до и после кризиса. Прогнозирование посткризисного периода на основе докризисных данных и последующее сравнение фактических и спрогнозированных значений, предположительно, является хорошим подходом к исследованиям такого рода.

Цели работы

Основной целью данной работы является выявление и анализ качественных изменений в экономике США, произошедших вследствие кризиса 2020 года. Для формальной и строгой постановки цели выдвинем следующую гипотезу:

H_0 : экономические показатели до и после кризиса 2020 года не имеют качественных различий.

Тогда альтернативная гипотеза будет выглядеть следующим образом:

H_1 : экономические показатели претерпели качественные изменения.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Собрать и систематизировать данные макроэкономических показателей США;
2. Сравнить качество моделей прогнозирования;
3. Спрогнозировать макроэкономические показатели второго периода на основе данных первого периода;
4. Сравнить фактические данные первого периода с фактическими данными второго периода и со спрогнозированными на второй период значениями.

Глава 1. Данные и теория

1.1 Описание показателей

Как было сказано выше, для данной работы были выбраны экономические данные США с 1947 года ввиду их доступности и целостности. Все данные были выгружены из онлайн-базы данных FRED, содержащей экономическую статистику из множества национальных, международных, государственных и частных источников. FRED создана и поддерживается исследовательским отделом Федерального резервного банка Сент-Луиса.

Ниже представлено описание каждого экономического показателя, задействованного в исследовании.

1. Реальный валовой внутренний продукт (rgdp) – это внутренний валовой продукт, измеренный в неизменных ценах базового года. Он представляет собой отношение номинального ВВП к индексу цен. Реальный ВВП позволяет исключить из номинального ВВП инфляцию или дефляцию и сравнить реальные объёмы производства в разные годы. ВВП измеряется как общая рыночная стоимость товаров и услуг, произведенных экономикой за определенный период.
2. Базовый индекс потребительских цен (core_cpi) – это показатель, который отражает изменение стоимости фиксированной потребительской корзины, исключая цены на продукты питания и энергоносители. Эти товары подвержены значительным колебаниям цен, которые могут исказить общую картину инфляции. Core CPI используется для более точного измерения базовой инфляции, поскольку он менее подвержен влиянию краткосрочных изменений цен на отдельные товары.
3. Реальная денежная масса M2 (m2) – общий объем денежных средств, доступных в экономике, скорректированный на инфляцию. В денежную массу M2 входят наличные деньги, дорожные чеки, сберегательные

- счета, депозиты, срочные вклады, ликвидные ценные бумаги, денежные фонды.
4. Реальный располагаемый личный доход (pers_inc) – это часть располагаемого дохода, которая идёт на потребление с учётом инфляции. Он позволяет оценить, сколько товаров и услуг можно купить на эти деньги. Располагаемый доход — это сумма дохода, которой домохозяйства действительно располагают и тратят по своему усмотрению: на личное потребление и личные сбережения. Он остаётся у домохозяйств после вычета из личного дохода индивидуальных налогов.
 5. Валовые частные внутренние инвестиции (gpdі) – общая сумма расходов на капитальные инвестиции частными предприятиями, домохозяйствами и некоммерческими организациями. GPDІ включает в себя затраты на физические активы, такие как оборудование, инструменты, здания и инфраструктура, которые используются для производства. Также в него входят расходы на строительство нового жилья.
 6. Уровень безработицы (unrate) – число безработных в процентах от рабочей силы. Сюда входят лица в возрасте 16 лет и старше в 50 штатах США.
 7. Индекс промышленного производства (indpro) – это показатель, который измеряет объём производства промышленного сектора страны, включая горнодобывающую промышленность, обрабатывающую промышленность и коммунальные услуги.
 8. Коммерческие и промышленные кредиты (loans) – это показатель, отражающий уровень коммерческого кредитования и использования капитала предприятиями.
 9. Новые жилые единицы (houst) – это индикатор, который показывает количество новых частных жилых единиц, начатых к строительству.

10. Ставка Федеральной резервной системы (`key_rate`) — это средняя процентная ставка, по которой банки США кредитуют друг друга на короткий срок (обычно на ночь), используя свои избыточные резервы. Этот показатель отражает состояние денежного рынка и является важным инструментом ФРС для контроля над денежной массой и инфляцией.

Далее для удобства вместо полного названия переменных будут использоваться сокращения, указанные в скобках.

1.2 Используемый инструментарий

В этом разделе указаны программные продукты и инструменты, использованные в ходе реализации проекта.

Основным языком программирования был выбран Python. Этот язык оптимально подходит для исследовательских проектов благодаря своей гибкости и наличию большого числа готовых библиотек, что значительно ускоряет процесс разработки.

Работа над проектом велась на платформе Google Colab, которая предоставляет бесплатный доступ к вычислительным ресурсам на базе CPU, что позволило эффективно выполнять вычислительные задачи и анализ данных. Были задействованы следующие библиотеки:

- Pandas – для работы с большими наборами данных;
- NumPy – для математических вычислений;
- Matplotlib – для визуализации данных, построения графиков;
- Statsmodels – для проведения статистических тестов и регрессионных моделей;
- SciPy – для статистических расчётов и вычислений;
- CatBoost и LightGBM – для одноимённых моделей градиентного бустинга;
- Fredapi – для взаимодействия с базой данных FRED.

1.3 Первичный анализ и обработка данных

Дадим определение временного ряда. Временной ряд – это последовательность наблюдений за некоторой величиной через равные промежутки времени на определённом временном отрезке.

Большинство переменных в данном исследовании представлены месячными данными. Однако одна из ключевых, *rgdp*, публикуется только ежеквартально. Для приведения остальных переменных к аналогичному временному масштабу использовался метод агрегирования "avg" из FRED API.

	<i>rgdp</i>	<i>core_cpi</i>	<i>m2</i>	<i>pers_inc</i>	<i>gpd</i>	<i>unrate</i>	<i>indpro</i>	<i>loans</i>	<i>houst</i>	<i>key_rate</i>
count	309	225	261	261	309	305	309	309	261	279
mean	10009,65	4,33	2940,09	8310,66	1271,06	5,7	59,6	707,11	1433,14	4,6
std	6084,86	2,69	1626,62	4345,85	1321,9	1,7	30,46	765,62	376,01	3,57
min	2172,43	0,7	992,8	2327,5	34,51	2,6	13,49	11,72	526	0,06
25%	4748,05	2,45	1848,3	4557,7	142,81	4,4	34,17	73,94	1212	1,89
50%	8604,22	3,35	2431,2	7299,3	835,83	5,5	54,71	487,3	1456	4,3
75%	15670,88	5,18	3542,3	12090,3	2075,79	6,8	91,93	1054,41	1655	6,27
max	22768,87	15,21	7651,1	18378,3	5006,85	13	103,98	2991,03	2424	17,78

Таблица 1. Сводная информация о данных

После первичного анализа основных статистических данных нашего набора мы обнаруживаем, что некоторые переменные имеют более 300 наблюдений, в то время как другие заметно меньше. Мы можем исключить переменные со слишком малым количеством наблюдений, но предпочтительнее будет их сохранить. Чтобы все столбцы имели одинаковую длину периода, мы удаляем начальные наблюдения, где имеются значения NaN.

Теперь у нас есть набор данных из 10 переменных, содержащий в общей сложности 225 ежеквартальных наблюдений, начиная с 1968 года.

	rgdp	core_cpi	m2	pers_inc	gpd	unrate	indpro	loans	houst	key_rate
1968-01-01	5057.553	3.822615	1550.7	3508.0	152.288	3.7	36.3840	86.5311	1455.0	4.79
1968-04-01	5142.033	4.263214	1561.0	3565.7	158.943	3.6	36.8745	89.4305	1456.0	5.98
1968-07-01	5181.859	4.882643	1570.1	3577.1	155.683	3.5	37.1606	91.2858	1521.0	5.95
1968-10-01	5202.212	5.429443	1586.8	3604.8	160.760	3.4	37.7249	94.5130	1582.0	5.92
1969-01-01	5283.597	5.873770	1594.4	3613.7	172.388	3.4	38.4588	97.9357	1678.0	6.57
...
2023-01-01	22112.329	6.467833	7002.7	16662.7	4725.828	3.5	102.5914	2808.6462	1369.0	4.52
2023-04-01	22225.350	5.942054	6854.4	16797.3	4780.290	3.6	102.7918	2789.8482	1455.0	4.99
2023-07-01	22490.692	5.179249	6782.5	16819.6	4915.033	3.7	103.2005	2746.8778	1380.0	5.26
2023-10-01	22679.255	4.708539	6724.1	16902.1	4954.426	3.7	102.6825	2754.9866	1481.0	5.33
2024-01-01	22768.866	4.505412	6682.4	16946.6	5006.854	3.8	102.3955	2761.3745	1403.0	5.33

Рисунок 1. Таблица с исходными данными

Следует визуально изучить временные ряды, чтобы иметь наглядное представление об их характеристиках.

В основном тексте работы для экономии места представлены графики только переменной unrate. Графики всех переменных вынесены в приложение.

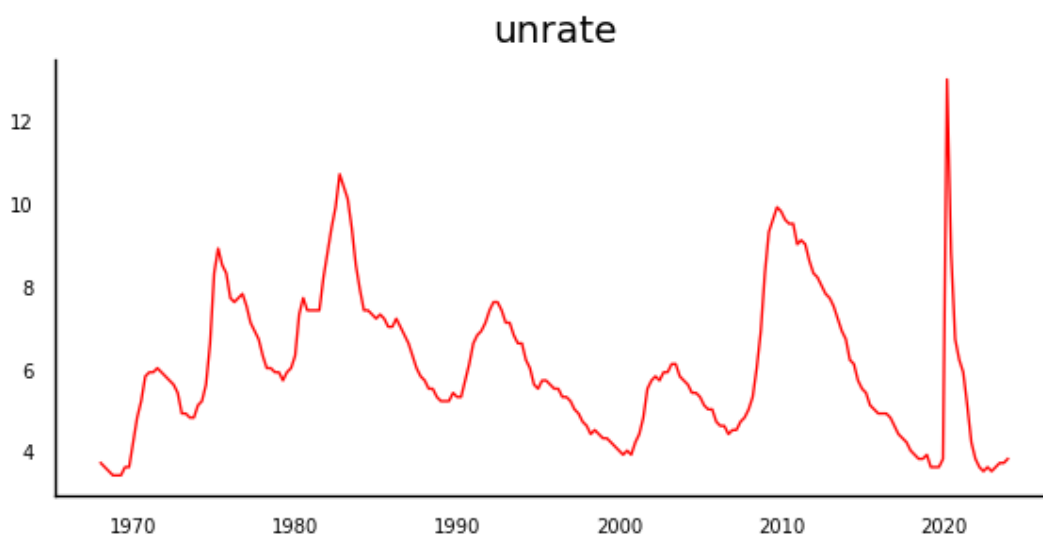


Рисунок 2. График безработицы

Можно видеть, что некоторые переменные, вроде m2 и pers_inc, находятся в явном восходящем тренде. Другие, в первую очередь unrate и houst, имеют некоторую сезонность. Также заметим, что общее поведение core_cpi и key значительно изменилось после 1980-х годов, что связано с

изменением монетарной политики США в 1970-х. Кроме того, все графики ярко отражают экономическое потрясение от пандемии Covid-19.

Проведём декомпозицию наших временных рядов. Это поможет выделить их составляющие: систематическую составляющую (тренд, сезонность) и случайную (нерегулярную) компоненту или остаток.

Функцию исходного ряда можно представить в следующем виде:

$$y(t) = x(t) + s(t) + z(t),$$

где $x(t)$ – тренд, устойчивая долговременная тенденция изменения значений временного ряда, закономерно изменяющаяся во времени;

$s(t)$ – сезонная составляющая, периодически повторяющаяся компонента временного ряда;

$z(t)$ – остаток – величина, показывающая нерегулярную (не описываемую трендом или сезонностью) составляющую исходного ряда в определённом временном интервале.

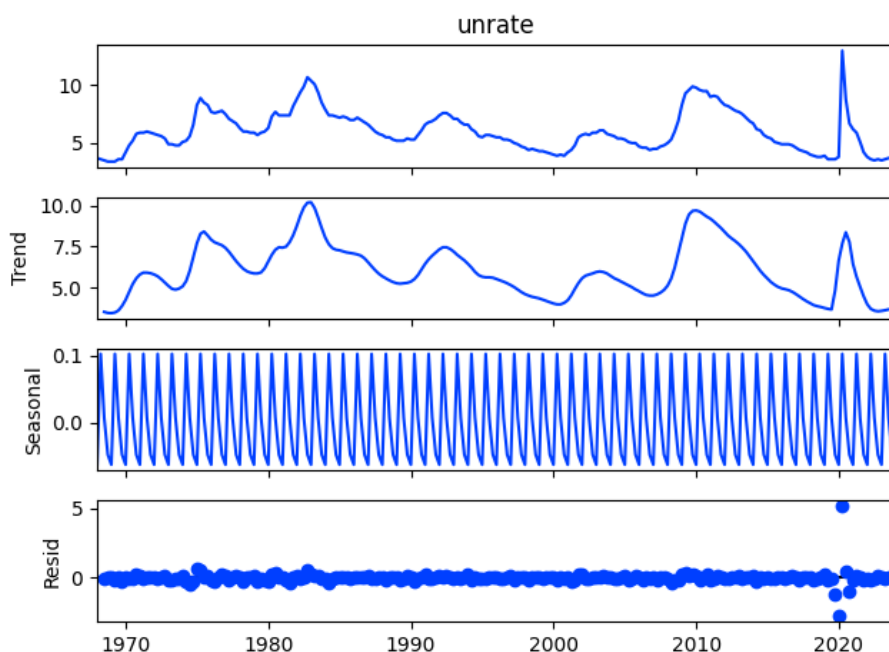


Рисунок 3. Декомпозиция временного ряда *unrate*

Теперь пора разделить имеющиеся данные на два периода и рассматривать их независимо друг от друга:

1. 01.01.1968 – 01.01.2020
2. 01.04.2021 – 01.01.2024

Таким образом мы исключаем данные за 4 квартала 2020 года.

1.4 Стационарность

В анализе временных рядов одной из ключевых концепций является стационарность. Стационарность временного ряда означает, что его статистические свойства, такие как среднее, дисперсия и автокорреляция, не изменяются во времени.

Свойство стационарности потребуется при дальнейшем анализе и прогнозировании. Модели, применённые в исследовании, для корректной работы требуют стационарности данных, чтобы избежать ложных выводов из-за изменяющихся статистических свойств. Хотя модели градиентного бустинга могут работать с нестационарными данными, результаты и точность прогнозов значительно улучшаются при использовании стационарных временных рядов.

По этим причинам следующим шагом следует изучить наши временные ряды на наличие этой важной характеристики. Для этого воспользуемся одной из самых распространённых проверок – тестом Дики-Фуллера (ADF-тест).

Основная идея теста заключается в подтверждении или отклонении двух гипотез:

- H_0 : временной ряд имеет единичный корень, то есть он нестационарный.
- H_1 : временной ряд не имеет единичного корня, то есть он является стационарным.

Для подтверждения или опровержения этих гипотез используются р-значения. Р-значение – это наименьшее значение уровня значимости, то есть вероятности отказа от справедливой гипотезы. В случае если р-значение меньше 5%, то мы отвергаем нулевую гипотезу, и, следовательно, можно с 95% вероятностью говорить, что ряд стационарен.

Сравнительная таблица будет представлена ниже для более удобного восприятия результатов. А перед этим приведём наши временные ряды к желаемому виду. Для этого существует множество методов. Начнём с самых простых.

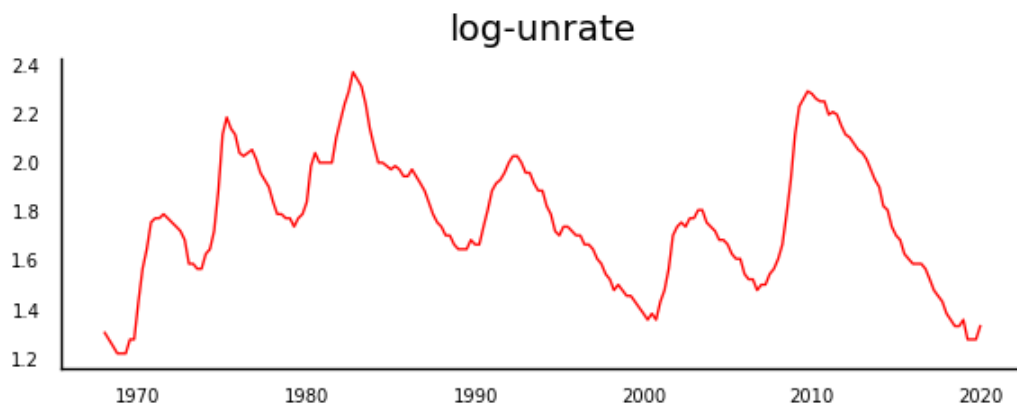


Рисунок 5. Логарифмированный ряд *unrate* доковидного периода

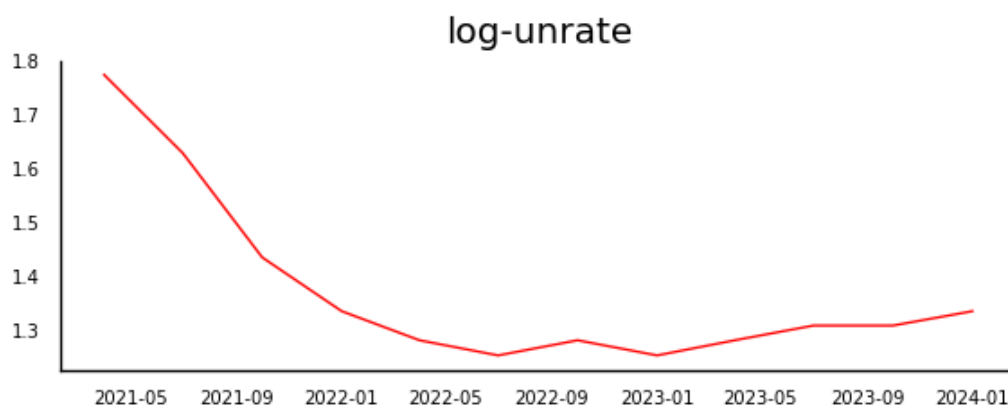


Рисунок 4. Логарифмированный ряд *unrate* постковидного периода

1. При помощи логарифмического преобразования данных мы разберёмся с проблемами масштаба. Можно отметить, что, например, когда у логарифмированного ряда наблюдается линейный тренд, это означает, что исходный ряд имеет экспоненциальный рост. На докризисном периоде это можно заметить у *m2*, *rgdp*, *pers_inc*, *gpdI* и *loans*. На посткризисном – у *rgdp*.
2. Следующее преобразование – дифференцирование. Если из нестационарного ряда первыми разностями удаётся получить стационарный, то он называется интегрированным рядом первого порядка.

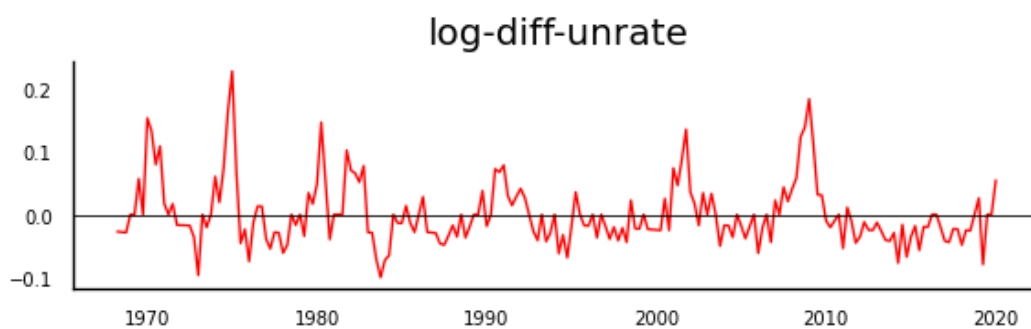


Рисунок 6. Преобразованный ряд *unrate* доковидного периода

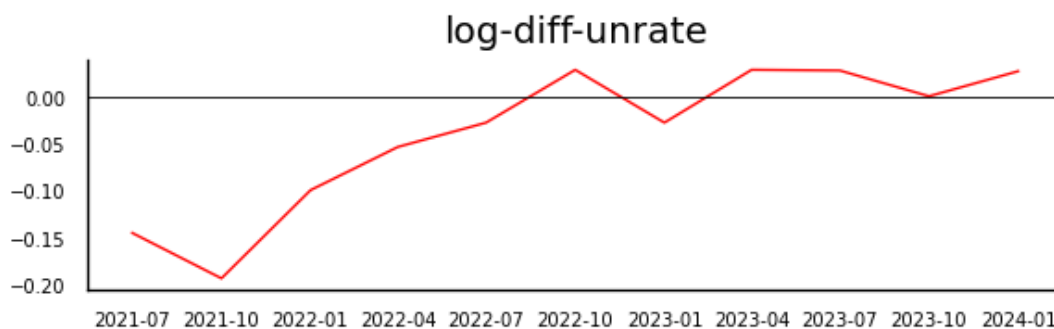


Рисунок 7. Преобразованный ряд *unrate* постковидного периода

Из сравнительной таблицы ниже видно, что применённых преобразований оказалось достаточно, чтобы получить стационарные временные ряды.

p-value				
Период	Доковидный		Постковидный	
Ряд	Исходный	Преобразованный	Исходный	Преобразованный
rqdp	0,997	$5,4 * 10^{-6}$	0,93	0,039
core_cpi	0,72	$0,1 * 10^{-7}$	0,087	0,029
m2	1	$1,6 * 10^{-6}$	$0,2 * 10^{-15}$	0,027
pers_inc	1	$1,9 * 10^{-9}$	0,735	0,016
gpdі	0,98	$2,3 * 10^{-20}$	0,01	0,036
unrate	0,046	$5,2 * 10^{-7}$	0,563	0,006
indpro	0,732	$0,1 * 10^{-4}$	0,004	0,007
loans	0,999	$0,7 * 10^{-5}$	0,419	0,041
houst	0,006	$1,7 * 10^{-18}$	$0,28 * 10^{-28}$	0,019
key_rate	0,422	$0,2 * 10^{-5}$	0,361	0

Таблица 2. Результаты преобразований

В доковидном периоде все временные ряды, кроме `unrate` и `houst`, имели нестационарный вид. Однако после преобразований можно отвергнуть нулевую гипотезу о нестационарности этих рядов, что позволяет считать их стационарными. В постковидном периоде временные ряды демонстрируют различный характер. Однако, после проведения соответствующих преобразований, нулевая гипотеза о нестационарности также отвергается. Дополнительно можно заметить, что с уменьшением объёма выборки p -значение у преобразованных данных растёт.

1.5 Автокорреляция

Будет полезно проанализировать автокорреляцию данных. Она поможет нам понять закономерности и зависимости между временными шагами. Одним из распространенных инструментов для этого анализа является функция автокорреляции (ACF). Графики ACF показывают корреляцию между временным рядом и его запаздывающими версиями.

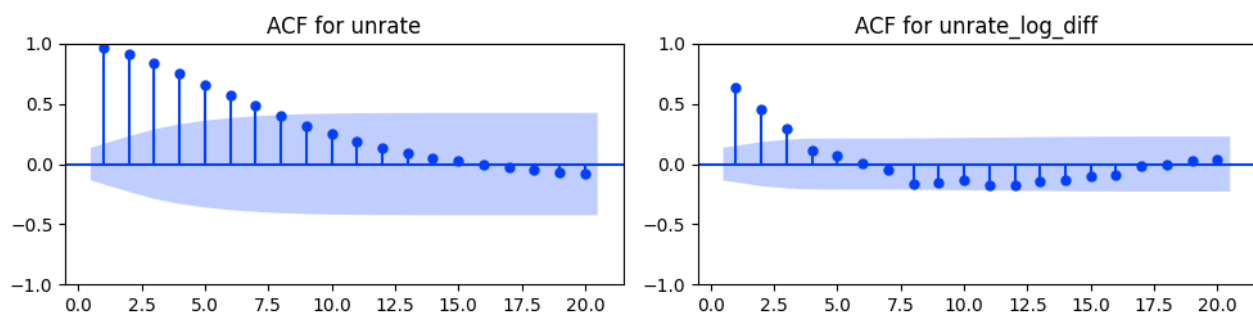


Рисунок 8. Графики автокорреляции ряда `unrate`

Медленно затухающий график ACF обычно указывает на наличие тренда в данных. После преобразований такого свойства уже не наблюдается, чем можно дополнительно подтвердить приведение наших рядов к стационарному виду.

Строить ACF для посткризисного периода нецелесообразно, поскольку малое количество имеющихся данных не позволяет быть уверенными в корректности результатов.

1.6 Матрицы корреляций

Построим матрицы корреляций для стационарных временных рядов на двух периодах. Если рассчитывать корреляцию на исходных рядах, результат может быть искажен из-за трендов, сезонных компонентов или других нестационарных свойств данных. Такие искажения могут привести к завышению корреляций, так как нестационарные ряды могут иметь общие тренды или сезонные компоненты, которые искусственно увеличивают связь между рядами.

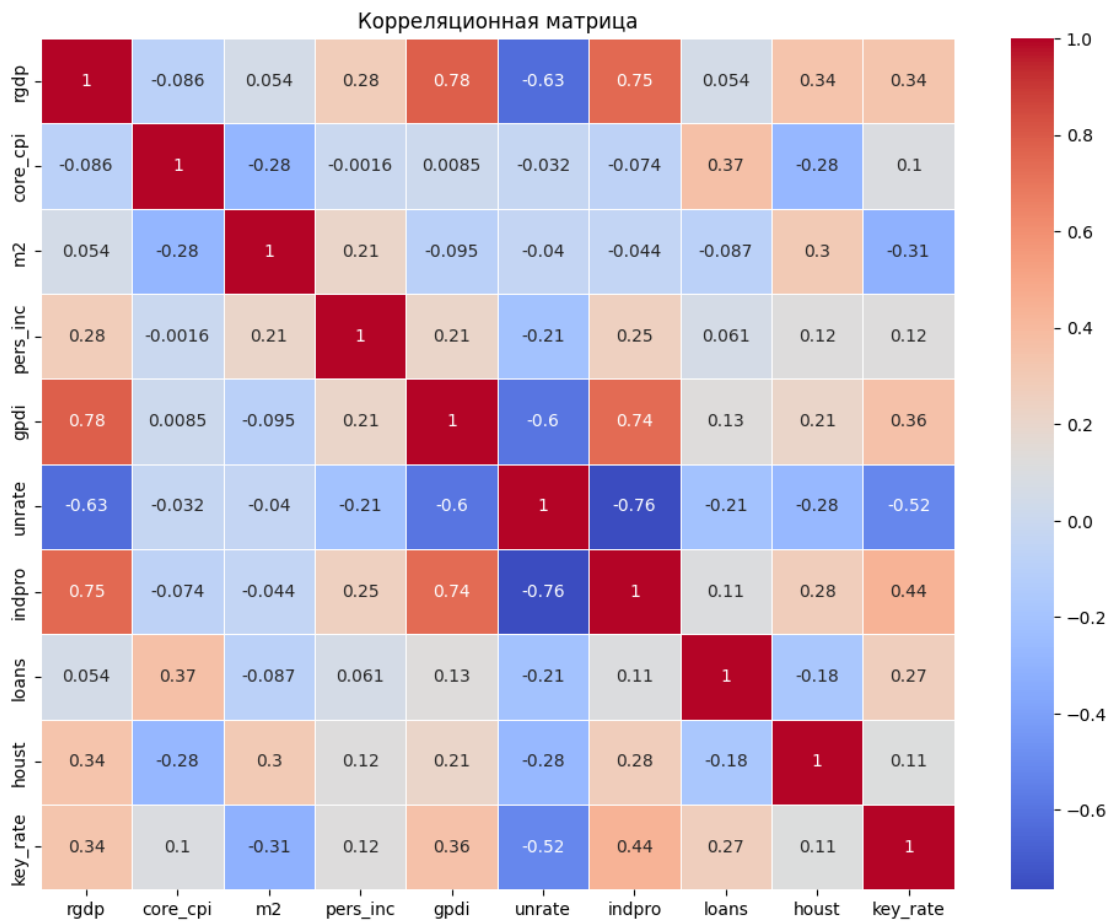


Рисунок 9. Корреляция переменных на доковидном периоде

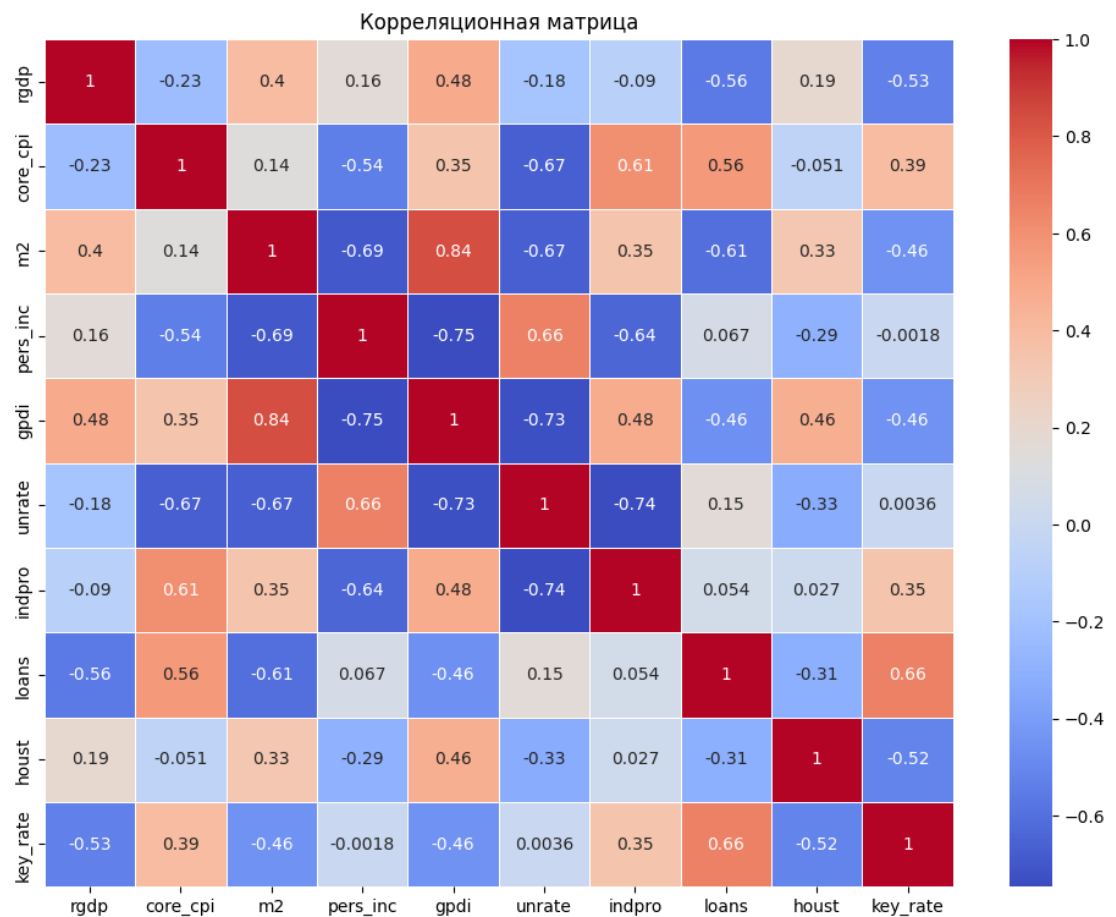


Рисунок 10. Корреляция переменных на постковидном периоде

Как видно из матриц, результаты демонстрируют значительную разнородность динамики. Некоторые пары переменных сменили отрицательную корреляцию на положительную, в то время как другие – наоборот. У одних пар корреляция усилилась, у других ослабла. Вероятной причиной такого поведения является недостаточное количество данных за посткризисный период, то есть мы имеем дело с незначимой выборкой. Так что это не корреляции, а случайные колебания.

1.7 Поправка Бонферрони

При проведении множественных статистических тестов одновременно возникает проблема увеличения вероятности ошибок первого рода, то есть вероятность отвергнуть нулевую гипотезу, когда она на самом деле верна. В таких случаях полезно использовать корректировки, которые помогают контролировать общий уровень значимости. Одной из наиболее популярных и простых корректировок является поправка Бонферрони.

Когда мы проводим несколько независимых тестов на одном и том же наборе данных, каждый тест имеет свою собственную вероятность ложного срабатывания (обозначаемую как α , уровень значимости). Если мы проводим m тестов с уровнем значимости α для каждого, то общая вероятность того, что хотя бы один из тестов даст ложноположительный результат, увеличивается и может быть приближенно оценена как $1 - (1 - \alpha)^m$. При увеличении количества тестов эта вероятность значительно возрастает, что может привести к ошибочным выводам.

В нашем исследовании мы планируем анализировать 10 временных рядов. Зафиксируем для уровня значимости стандартное значение $\alpha = 0,05$ и посчитаем новый уровень значимости:

$$\alpha_{new} = \frac{\alpha}{m} = \frac{0,05}{10} = 0,005$$

Таким образом, согласно поправке Бонферрони, во всех дальнейших тестах будет использоваться уровень значимости $\alpha = 0,005$.

1.8 Тест Гренджера

Идея теста заключается в том, что значения (изменения) временного ряда, являющегося причиной изменений некоторого другого временного ряда, должны предшествовать изменениям этого временного ряда, и кроме того, должны вносить значимый вклад в прогноз его значений.

В тесте Гренджера последовательно проверяются две нулевые гипотезы: « x не является причиной y по Гренджеру» и « y не является причиной x по Гренджеру». Для проверки этих гипотез строятся две регрессии: в каждой регрессии зависимой переменной является одна из проверяемых на причинность переменных, а регрессорами выступают лаги обеих переменных.

Причинность в смысле Гренджера не обязательно означает истинную причинно-следственную связь. Это всего лишь статистическая взаимосвязь, которая указывает на то, что изменения в одном временном ряду предшествуют и помогают предсказывать изменения в другом.

Для теста Гренджера следует отдельно посчитать уровень значимости с поправкой Бонферрони:

$$\alpha_{granger} = \frac{0,05}{90} = 0,00056$$

Мы будем действовать по аналогичному ADF-тесту правилу. Если p -значение для пары переменных окажется меньше $\alpha_{granger}$, то можно говорить, что переменная «причины» оказывает ощутимое влияние на будущие значения «следствия».

Далее представлены результаты теста Гренджера с описанием. В таблице в заголовках столбцов к переменным приписан « x » – это переменные «причины». В индексах строк к переменным приписан « y » – это переменные «следствия».

	rgdp_x	core_cpi_x	m2_x	pers_inc_x	gpdi_x	unrate_x	indpro_x	loans_x	houst_x	key_rate_x
rgdp_y	1	0,0001	0,0003	0,0058	0,0845	0,0001	0	0,0745	0	0,1845
core_cpi_y	0	1	0,0011	0,0523	0	0	0	0,0008	0,0088	0
m2_y	0,0299	0,9538	1	0,0719	0,0665	0,0009	0,0003	0,332	0,2806	0,0104
pers_inc_y	0,0001	0,1282	0,0569	1	0,0005	0,0087	0,0001	0,758	0,0203	0,6156
gpdi_y	0	0,0003	0,0549	0,011	1	0	0	0,0059	0	0,0008
unrate_y	0,0219	0,0001	0,001	0,0939	0,7695	1	0,0022	0,0085	0	0,4365
indpro_y	0,035	0,0001	0,0009	0,2731	0,0022	0,007	1	0,0004	0	0,0625
loans_y	0	0,0065	0,3294	0,0033	0	0	0	1	0,0009	0
houst_y	0,0031	0,1566	0,0121	0,5568	0,0065	0,0001	0,0001	0,0006	1	0,0192
key_rate_y	0,0001	0,1159	0,0533	0,0133	0,008	0	0,0024	0,3271	0,0027	1

Таблица 3. Результаты теста Гренджера на доковидном периоде

	rgdp_x	core_cpi_x	m2_x	pers_inc_x	gpdi_x	unrate_x	indpro_x	loans_x	houst_x	key_rate_x
rgdp_y	1	0,0017	0,8105	0,3476	0,3149	0,2716	0,6437	0,0021	0,5174	0,7535
core_cpi_y	0,3882	1	0	0,0117	0,0185	0	0,0515	0,001	0,5803	0,9356
m2_y	0,2346	0	1	0,0919	0,3234	0,1301	0,6023	0,0011	0,3245	0,5897
pers_inc_y	0,0176	0,9556	0,002	1	0,0061	0,0004	0,1081	0,0544	0,1148	0,0868
gpdi_y	0,4074	0,1533	0,0317	0,3582	1	0,2991	0,6452	0,0002	0,7995	0,716
unrate_y	0,8858	0,1117	0,0426	0,1971	0,1593	1	0,7448	0,0224	0,894	0,811
indpro_y	0,5417	0,5254	0,2763	0,1063	0,1436	0,0162	1	0,5059	0,0043	0,2988
loans_y	0,9221	0	0	0	0	0,0001	0,0045	1	0,8025	0,9198
houst_y	0,0213	0,2675	0,1082	0,7321	0,1384	0,5159	0,8915	0,0002	1	0
key_rate_y	0,0087	0,0015	0,5542	0,0068	0,2541	0,218	0,114	0,0241	0,4078	1

Таблица 4. Результаты теста Гренджера на постковидном периоде

Чем больше выборка, тем более весомым является подтверждение нулевой гипотезы. В постковидном периоде выборка небольшая, поэтому большие р-значения лучше интерпретировать не как отсутствие связи, а как неуспех её обнаружения. Оказалось, что у многих переменных связь не подтвердилась. Список таких пар вынесен в приложение.

Хорошей практикой при прогнозировании какого-либо конкретного показателя является сохранение только тех переменных, которые обладают наибольшей объяснительной силой для него, и исключение тех, которые не являются особо значимыми. Однако, поскольку нас больше интересует исследование всего набора макроэкономических переменных, исключение любых из них приведёт к ухудшению результатов прогнозирования регрессионными моделями.

Глава 2. Прогнозирование и исследование

2.1 Модели прогнозирования

Давайте посмотрим на модели, которые будут использованы для прогнозирования:

1. VAR (Vector Autoregression): Модель векторной авторегрессии используется для моделирования и прогнозирования взаимозависимых временных рядов. VAR учитывает линейные взаимосвязи между множественными временными рядами, моделируя каждый временной ряд как функцию прошлых значений всех временных рядов в системе.
2. VECM (Vector Error Correction Model): Модель векторной коррекции ошибок используется для временных рядов, которые имеют долгосрочные равновесные отношения (коинтеграция). VECM включает в себя компоненты коррекции ошибок, которые корректируют краткосрочные отклонения от долгосрочного равновесия.
3. VARMA (Vector Autoregressive Moving Average): Модель векторной авторегрессии-скользящего среднего сочетает в себе элементы VAR и скользящего среднего (MA) для моделирования временных рядов. VARMA модель учитывает как автокорреляцию, так и зависимость текущих значений временных рядов от прошлых ошибок (шумов).
4. LightGBM (Light Gradient Boosting Machine): LightGBM — это алгоритм градиентного бустинга, разработанный с упором на высокую скорость и эффективность.
5. CatBoost (Categorical Boosting): CatBoost — это алгоритм градиентного бустинга, который использует технику одновременного построения деревьев для уменьшения переобучения. CatBoost известен своей стабильностью и высокой точностью.

Векторно-регрессионные модели VAR, VECM, VARMA выбраны в первую очередь потому, что в данной работе важным аспектом как раз являются взаимосвязи макроэкономических показателей между собой.

Чтобы получить более сбалансированное и полное представление о точности прогнозов, мы дополним список уже упомянутыми моделями градиентного бустинга.

2.2 Ошибки

При прогнозировании временных рядов важно оценивать точность предсказаний. Для этого используются различные метрики, такие как MAPE, MAE и RMSE. Ниже приведены краткие описания каждой из этих метрик:

1. MAPE (Mean Absolute Percentage Error):

MAPE представляет собой среднюю абсолютную процентную ошибку и рассчитывается как среднее значение абсолютных ошибок в процентах от фактических значений. Она позволяет интерпретировать ошибку прогноза в виде процента, что делает её удобной для сравнения различных временных рядов.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|,$$

где y_t – фактические значения, \hat{y}_t – прогнозируемые значения, n – количество наблюдений.

Примечание: далее вместо MAPE будет использоваться метрика "Точность", вычисляемая как $1 - MAPE$.

2. MAE (Mean Absolute Error):

MAE представляет собой среднюю абсолютную ошибку и рассчитывается как среднее значение абсолютных различий между фактическими и прогнозируемыми значениями. Эта метрика проста для интерпретации и позволяет понять, насколько в среднем прогнозируемые значения отклоняются от фактических.

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

3. RMSE (Root Mean Squared Error):

RMSE представляет собой корень из средней квадратической ошибки и рассчитывается как квадратный корень из среднего значения квадратов различий между фактическими и прогнозируемыми значениями. Эта метрика чувствительна к большим ошибкам, так как квадратичное возведение увеличивает влияние крупных отклонений.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

Использование этих метрик позволяет объективно оценить качество прогнозов. Важно отметить, что здесь и далее при вычислении ошибок в расчёт не берутся первые 4 спрогнозированных квартала. Это сделано для того, чтобы избежать неверной интерпретации результатов из-за сильных колебаний и волатильности фактических временных рядов в период 2020 года.

2.3 Тестирование моделей

Итак, перейдем к исследованию разницы между до- и посткризисным периодами. Нам необходимо спрогнозировать временные ряды на 16 кварталов (01.04.2020 – 01.01.2024). Первым делом посмотрим, как ведут себя модели на доковидном периоде. Основной период прогноза – 16 кварталов, однако, чтобы получить более точное представление о работе моделей, в качестве дополнительного периода для тестирования возьмём 8 кварталов. Для обучения моделей будем использовать все оставшиеся данные, предшествующие периоду прогноза.

I. 8 кварталов

1. Данные для обучения: 01.04.1968 – 01.01.2018
2. Период прогнозирования: 01.04.2018 – 01.01.2020

II. 16 кварталов

1. Данные для обучения: 01.04.1968 – 01.01.2016

2. Период прогнозирования: 01.04.2016 – 01.01.2020

Модель	VAR	VECM	VARMA	LGBM	CatBoost
На 8 кварталов					
Точность	0,943	0,935	0,952	0,947	0,913
MAE	85	89	64	49	51
RMSE	150	137	101	97	78
На 16 кварталов					
Точность	0,871	0,9	0,863	0,871	0,877
MAE	92	184	153	49	68
RMSE	150	342	232	86	114

Таблица 5. Результаты тестирования моделей

Точность всех моделей при прогнозировании на 8 кварталов примерно одинакова. При прогнозировании на 16 кварталов точность, как и ожидалось, снижается. Однако можно заметить, что для регрессионных моделей значения ошибок MAE и RMSE значительно увеличиваются, тогда как для бустинговых моделей такого не наблюдается.

Тем не менее мы не станем отказываться от какой-либо модели, так как, имея больше различных прогнозов, мы в дальнейшем сможем делать более верные выводы при анализе.

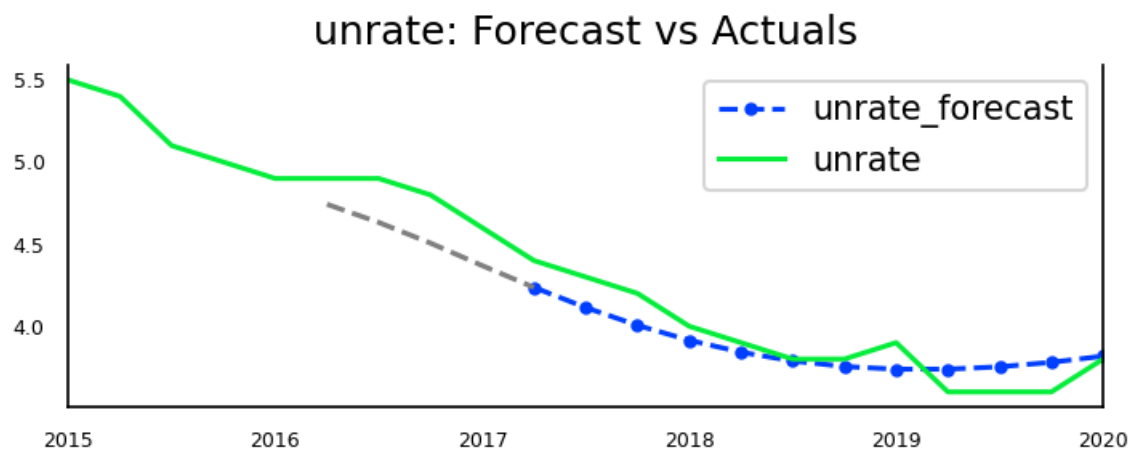


Рисунок 11. Прогноз ряда unrate моделью VECM

2.4 Прогнозирование на посткризисный период

Теперь можно спрогнозировать наши временные ряды на период с 2020 года (на 16 кварталов):

1. Данные для обучения: 01.04.1968 – 01.01.2020
2. Период прогнозирования: 01.04.2020 – 01.01.2024

Модель	VAR	VECM	VARMA	LGBM	CatBoost
Точность	0,488	0,292	0,536	0,409	0,422
MAE	213	228	196	200	192
RMSE	356	359	332	349	347

Таблица 6. Результаты моделей на посткризисном периоде

Из таблицы мы видим серьёзное ухудшение качества прогнозов на постковидный период по сравнению с тестированием на доковидном периоде. На данном этапе можно предположить, что наши макроэкономические переменные претерпели некоторые структурные сдвиги после 2020 года. Однако мы вернёмся к этим результатам позже.

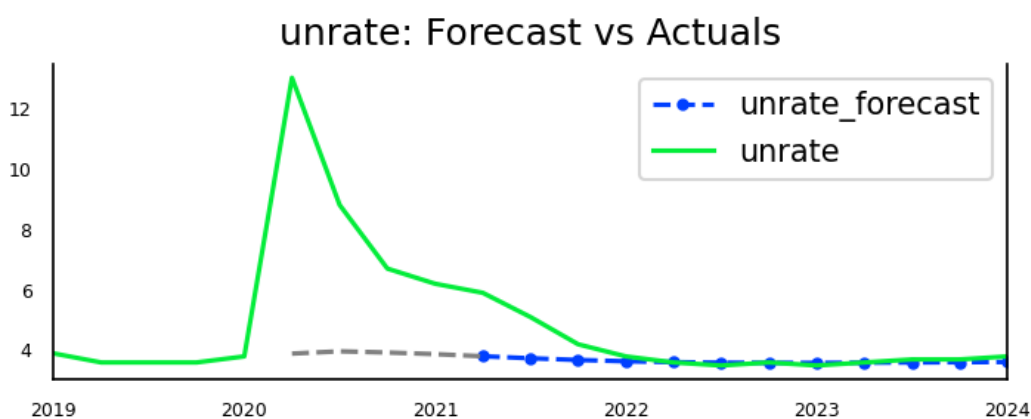


Рисунок 12. Прогнозирование временного ряда unrate на посткризисный период моделью VAR

2.5 Нормальное распределение

Проведём тест Шапиро-Уилка. Этот тест используется для оценки того, можно ли считать выборку из некоторой генеральной совокупности нормально распределённой. Многие статистические методы требуют, чтобы данные были нормально распределены. Если данные не соответствуют этому

предположению, использование таких методов может привести к неверным выводам.

Тест Шапиро-Уилка основан на гипотезе о том, что данные были взяты из нормально распределённой популяции. Он вычисляет статистику W , основанную на различиях между наблюдаемыми и ожидаемыми значениями в нормальном распределении. Значение W близкое к 1 указывает на то, что данные имеют нормальное распределение. В случае отклонения от нормальности распределения p -значение теста Шапиро-Уилка становится маленьким.

p-value		
Период	Доковидный	Постковидный
rqdp	0.00000852	0.08959
core_cpi	0.00000923	0.11186
m2	0.00008790	0.20297
pers_inc	$< 10^{-8}$	0.02831
gpd	0.00001125	0.01970
unrate	$< 10^{-8}$	0.00014
indpro	$< 10^{-8}$	0.00057
loans	0.00036416	0.45278
houst	0.00393847	0.1374
key_rate	$< 10^{-8}$	0.00052

Таблица 7. Результаты теста Шапиро-Уилка

Экстремально низкие p -значения в большинстве случаев, говорят нам о том, что далее следует учитывать ненормальность распределения данных.

2.7 U-критерий Манна-Уитни

Теперь проведём сравнение временных рядов с использованием критерия Манна-Уитни. Этот критерий является непараметрическим статистическим тестом, предназначенным для сравнения двух независимых выборок. Свойство непараметричности теста означает, что он не требует предположений о нормальном распределении данных.

Формулировка гипотез:

- H_0 : Распределения спрогнозированного и фактического временных рядов идентичны;
- H_1 : Распределения рядов различны.

Рассмотрим принцип работы критерия Манна-Уитни.

1. Пусть x_1, x_2, \dots, x_n – фактические значения, а y_1, y_2, \dots, y_m – спрогнозированные значения.
2. Данные объединяются в один массив: $Z = \{x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m\}$
3. Данные ранжируются. Ранг – это порядок, в котором элементы данных располагаются по величине. Присвоение рангов элементам – это процесс их сортировки и назначения порядковых номеров, которые отражают их положение в отсортированном массиве. Если есть совпадающие значения, используется средний ранг этих элементов.
4. Вычисляются суммы рангов для каждой группы:
 - Сумма рангов для фактических значений: $R_x = \sum_{i=1}^n \text{ранг}(x_i)$
 - Сумма рангов для спрогнозированных значений: $R_y = \sum_{j=1}^m \text{ранг}(y_j)$
5. Вычисляется U-статистика:
 - $U_x = nm + \frac{n(n+1)}{2} - R_x$
 - $U_y = nm + \frac{m(m+1)}{2} - R_y$
6. Используется наименьшее из значений U_x и U_y как статистика теста:
 $U = \min(U_x, U_y)$
7. Если p-значение для U меньше заданного уровня значимости, то нулевая гипотеза отвергается, что указывает на статистически значимые различия временного ряда.

	VAR	VECM	VARMA	LGBM	CatBoost
rgdp	+	+	+	+	+
core_cpi	+	+	+	+	+
m2					
pers_inc	+	+	+	+	+
gdpi	+	+	+	+	+
unrate	+	+	+	+	+
indpro	+	+	+	+	+
loans	+	+	+	+	+
houst	+	+	+	+	+
key_rate	+	+			

Таблица 8. Результаты теста Манна-Уитни

В случае, если в таблице указан знак "+", это означает, что тест на данных конкретной модели не обнаружил достаточных оснований для того, чтобы утверждать значимое различие между конкретным спрогнозированным временным рядом и его фактическими значениями. Другими словами, р-значение оказалось больше заданного ранее уровня значимости $\alpha = 0,005$ и нулевая гипотеза не была отвергнута.

Согласно результатам теста, две переменные, а именно m2 и key_rate, демонстрируют статистически значимые различия. Оценим их визуально.

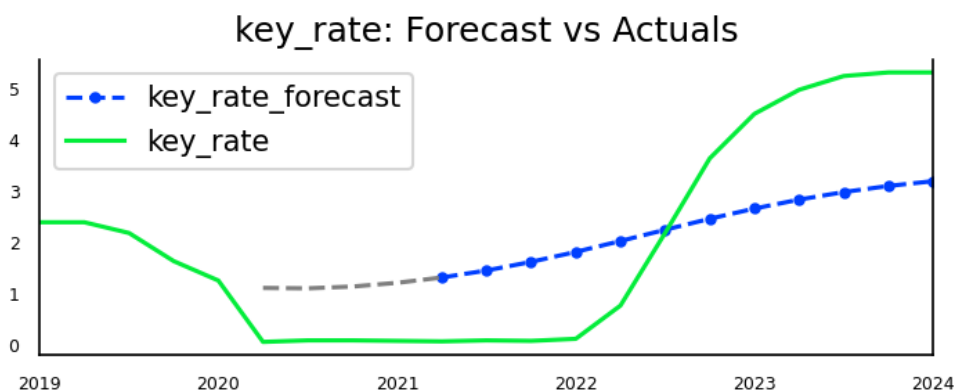


Рисунок 13. Прогноз key_rate моделью VECM

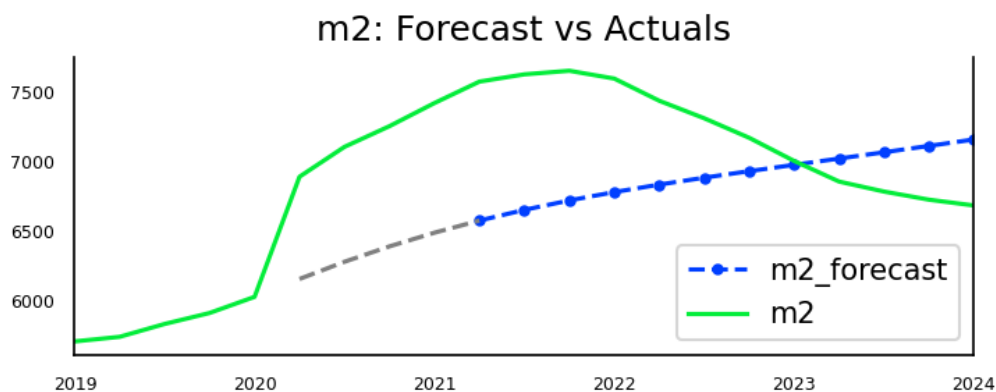


Рисунок 14. Прогноз m2 моделью VECM

Действительно, наблюдается значительное расхождение между фактическими и спрогнозированными значениями этих переменных в том числе и после ковидного шока. В связи с этим целесообразно проверить точность моделей, исключив из расчётов ошибок эти переменные с выявленными отклонениями.

Модель	VAR	VECM	VARMA	LGBM	CatBoost
Точность	0,884	0,859	0,875	0,911	0,914
MAE	194	220	173	171	158
RMSE	308	337	293	282	267

Таблица 9. Результаты моделей без выделяющихся переменных

Несмотря на то, что ошибки MAE и RMSE остаются сравнительно большими даже без учета двух указанных переменных, точность моделей заметно улучшается. В таком случае мы можем считать, что экономические показатели не претерпели значимых структурных изменений после ковидного кризиса.

Здесь же следует отметить, что модели градиентного бустинга, как видно из результатов прогнозирования на различных периодах, не превосходят по качеству векторно-регрессионные модели при прогнозировании взаимозависимых временных рядов.

Прогнозирование на будущее

В качестве небольшого дополнения, давайте выполним прогнозирование на предстоящие восемь кварталов.

Модель	VAR	VECM	VARMA	LGBM	CatBoost
Точность	0,811	0,81	0,819	0,835	0,813
MAE	299	229	238	198	300
RMSE	488	369	446	391	499

Таблица 10. Результаты тестирования моделей на всех известных данных

При оценке ожидаемого качества прогнозов на будущее следует учитывать результаты тестирования как при прогнозировании на последние известные 8 кварталов, так и при тестировании на 8 кварталах доковидного периода. Поскольку несмотря на то, что модели обучались на более длительном временном отрезке в первом случае, они, вероятно, были «сбиты с толку» непредсказуемыми всплесками данных за кризисный год, как это хорошо заметно по таблице результатов. Далее же, последние 8 кварталов включены в диапазон для обучения, что должно положительно повлиять на качество прогноза.

Наконец, давайте спрогнозируем наши макроэкономические показатели на ближайшие 2 года:

1. Данные для обучения: 01.04.1968 – 01.01.2024
2. Период прогнозирования: 01.04.2024 – 01.01.2026

На рисунке ниже графики построены не для всех моделей, чтобы он не выглядел слишком нагруженным.

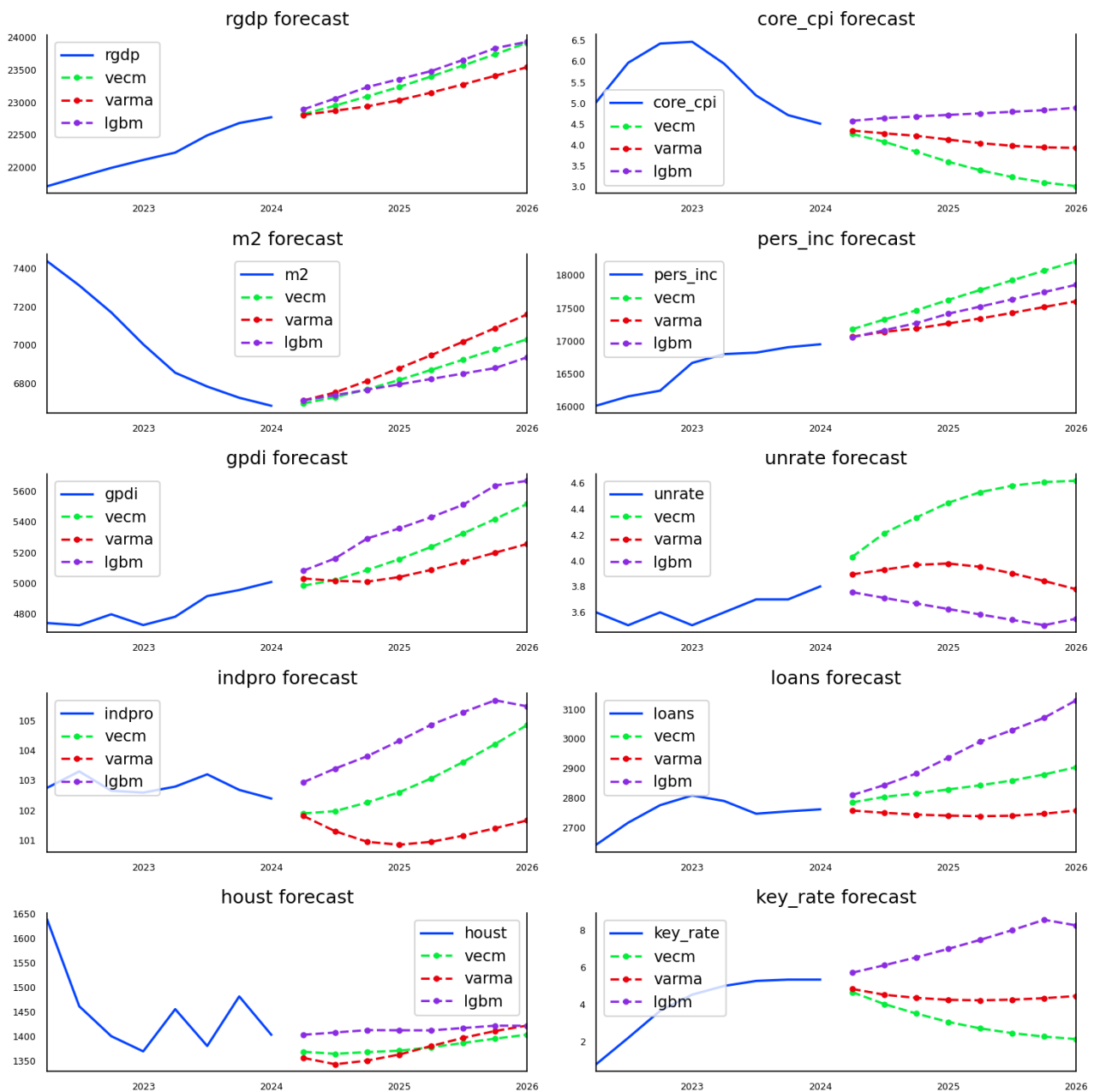


Рисунок 15. Прогноз на будущее

Заключение

В данной работе были исследованы макроэкономические данные США с целью выяснения, претерпели ли временные ряды, отражающие состояние экономики, качественные изменения после кризиса, вызванного пандемией Covid-19. Первоначальный анализ с использованием простых методов, таких как матрица корреляций и тест Гренджера, не позволил сделать практически значимых выводов.

Для более глубокого понимания ситуации мы перешли к прогнозированию временных рядов, основываясь на данных до 2020 года. Используя критерий Манна-Уитни, было выявлено, что на посткризисном периоде две переменные, $m2$ (денежная масса) и key_rate (ключевая ставка), демонстрируют статистически значимые различия между фактическими и спрогнозированными значениями. Убрав их из расчёта точности моделей, мы подтвердили их качественные изменения в посткризисном периоде.

С экономической точки зрения, значительное расхождение $m2$ и key_rate с реальными данными объясняется следующими факторами. В условиях паники и экономической дестабилизации ФРС США резко снизила ставки практически до нуля и начала активно печатать деньги, чтобы сгладить последствия экономического кризиса. Это привело к резкому росту объёма денежной массы в системе, что, в свою очередь, вызвало кратный прирост инфляции. В ответ на это ФРС была вынуждена повысить ключевую ставку выше её прежнего уровня. Эти меры были приняты быстро и в условиях высокой неопределённости, что делает их практически непредсказуемыми для любой модели прогнозирования. Конечно, это поверхностное объяснение, но оно отражает основную картину.

Таким образом, основной и наиболее важный вывод нашего исследования заключается в том, что макроэкономические показатели США не претерпели значимых структурных изменений после кризиса 2020 года, вызванного пандемией Covid-19. Иначе говоря, мы не отвергаем нулевую

гипотезу, сформулированную в качестве основного вопроса исследования. Несмотря на временные отклонения в отдельных показателях, экономика США продемонстрировала тенденцию к стабилизации и возвращению к своим нормальным значениям.

Список использованной литературы

1. Артамонов Н. В., Ивин Е. А., Курбацкий А. Н., Фантащини Д. Введение в анализ временных рядов: учебное пособие для вузов – Вологда: ВолНЦ РАН – 2021;
2. Афанасьев В.Н. Анализ временных рядов и прогнозирование: учебник / Афанасьев В.Н. — Саратов: Ай Пи Ар Медиа, 2020. — 310 с.;
3. Кизбикенов, К. О. Прогнозирование и временные ряды: учебное пособие / К. О. Кизбикенов. — Барнаул: АлтГПУ, 2017. — 115 с.;
4. Ковалевский, А. П. Статистические критерии: учебное пособие / Новосибирский государственный университет. – Новосибирск, 2022. – 92 с.;
5. Магнус Я. Р., Катышев П. К., Пересецкий А. А. Эконометрика. Начальный курс: Учеб. — 9-е изд., испр. — 2021. — 504 с.;
6. Christos Katris "Prediction of Unemployment Rates with Time Series and Machine Learning Techniques," Computational Economics, Springer; Society for Computational Economics – 2020;
7. Mann H. B., Whitney D. R. On a test of whether one of two random variables is stochastically larger than the other. // Annals of Mathematical Statistics. — 1947;
8. FRED. *Fed Reserve Economic Research*. Экономические данные ФРС: <https://fred.stlouisfed.org/>.

Приложение

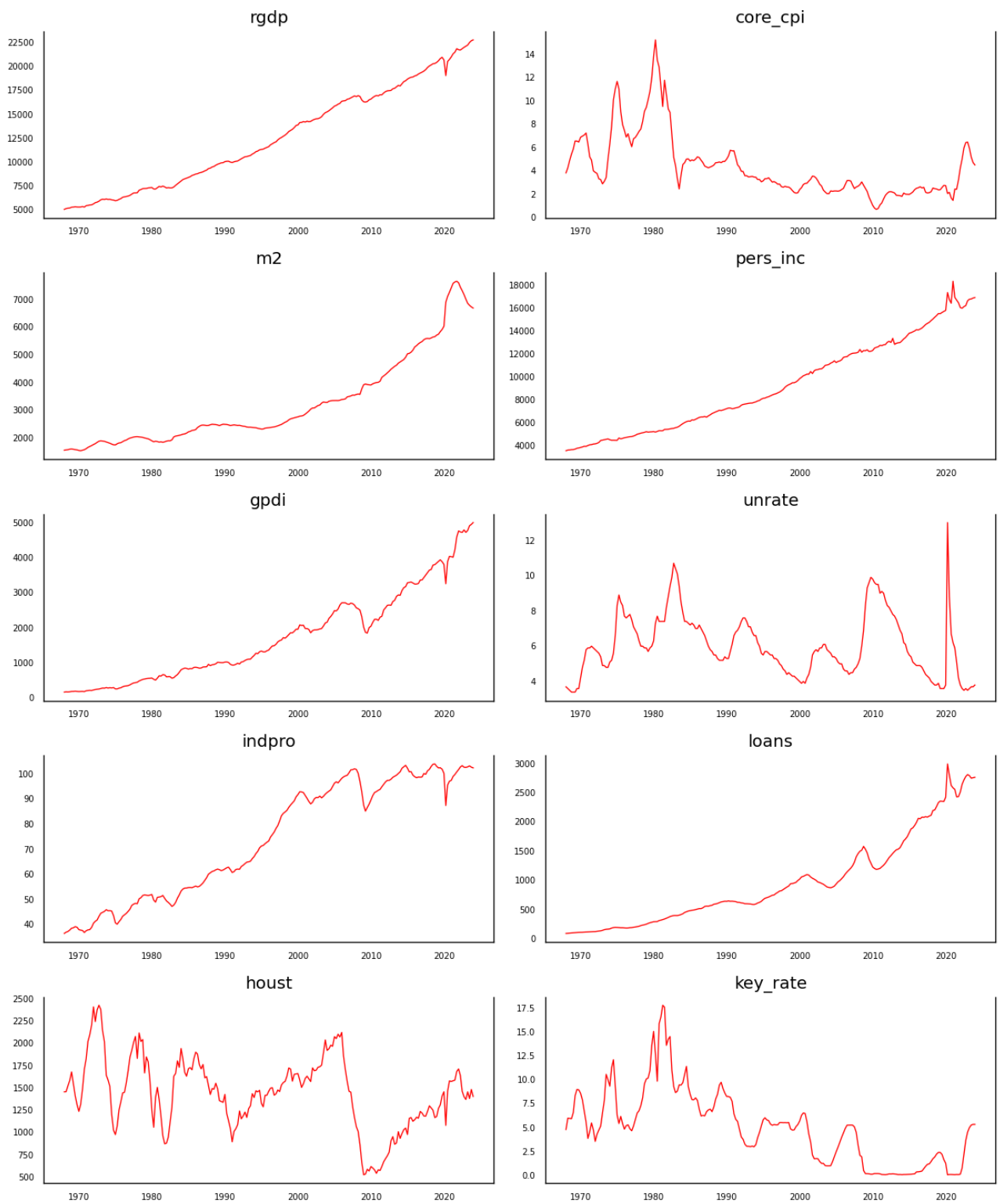


Рисунок 16. Графики исходных данных

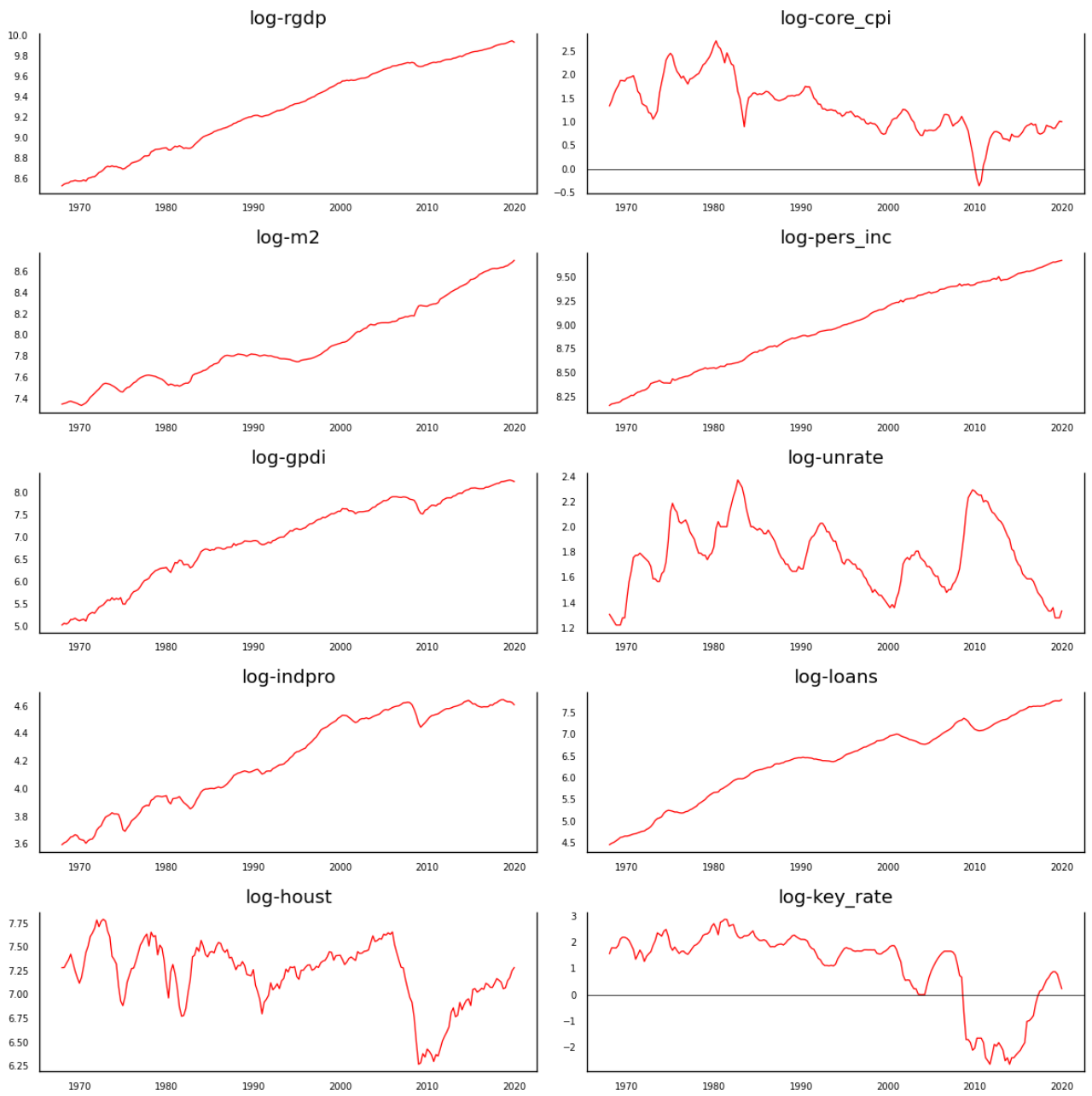


Рисунок 17. Графики логарифмированных данных доковидного периода

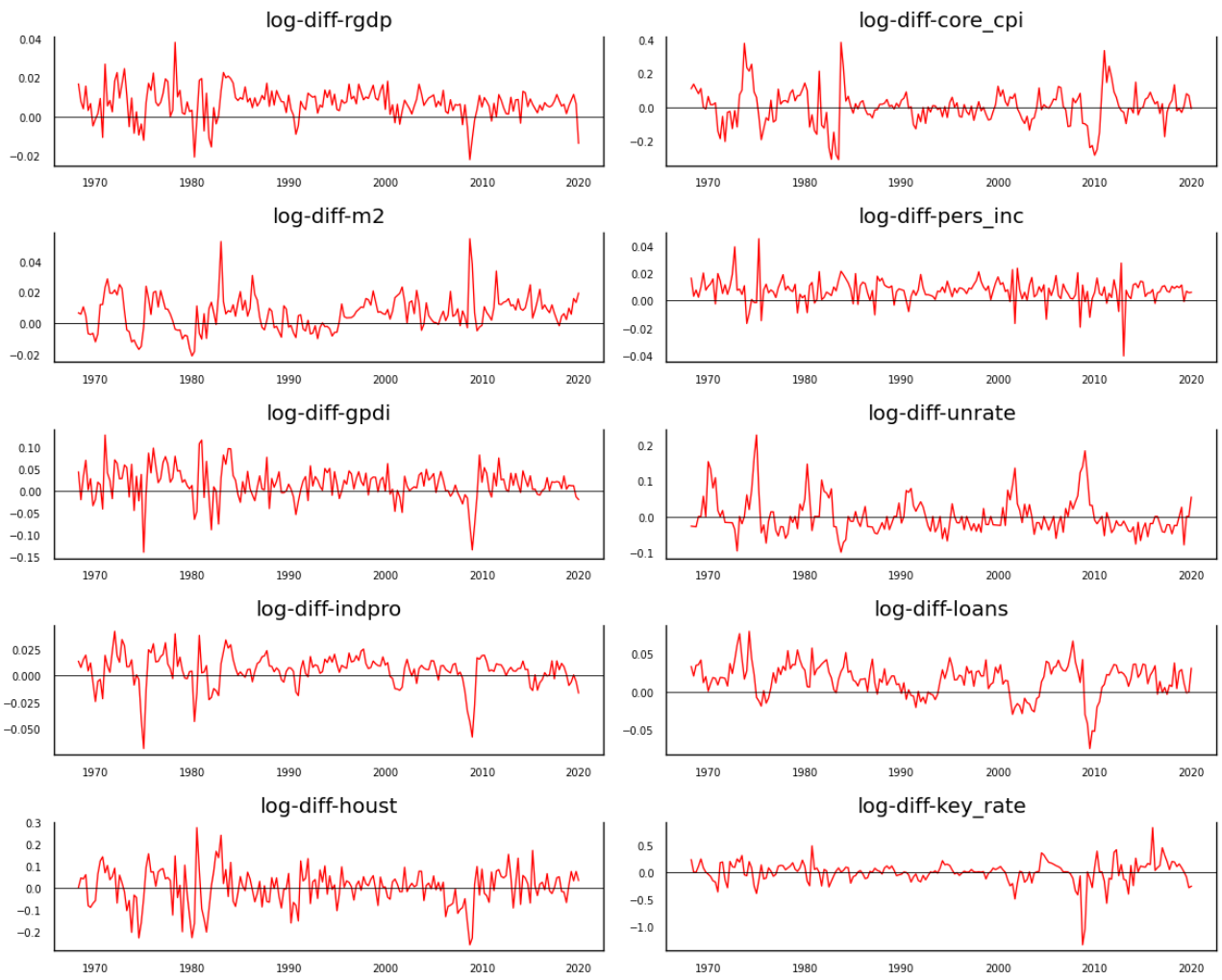


Рисунок 18. Графики преобразованных данных доковидного периода

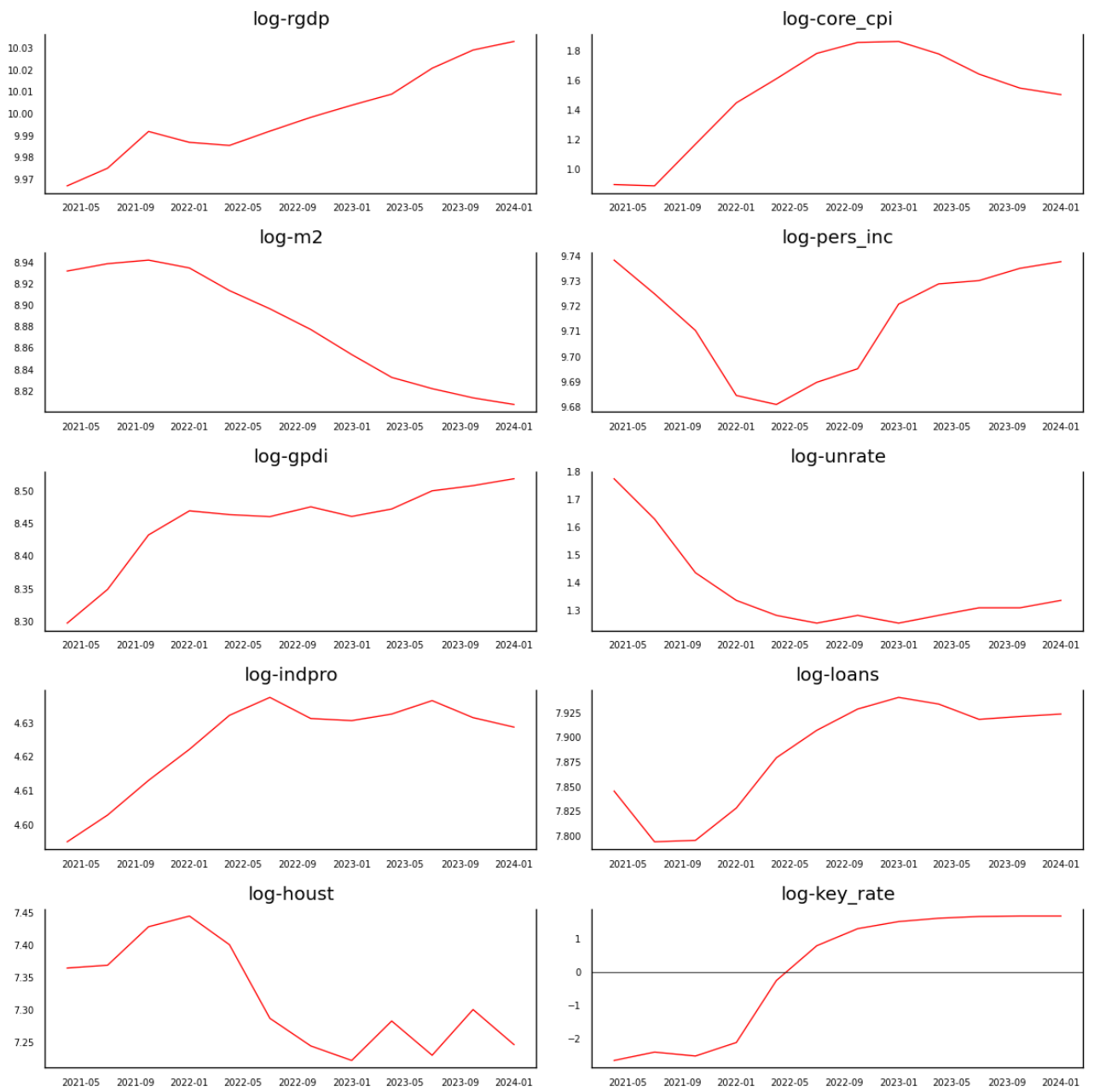


Рисунок 19. Графики логарифмированных данных постковидного периода

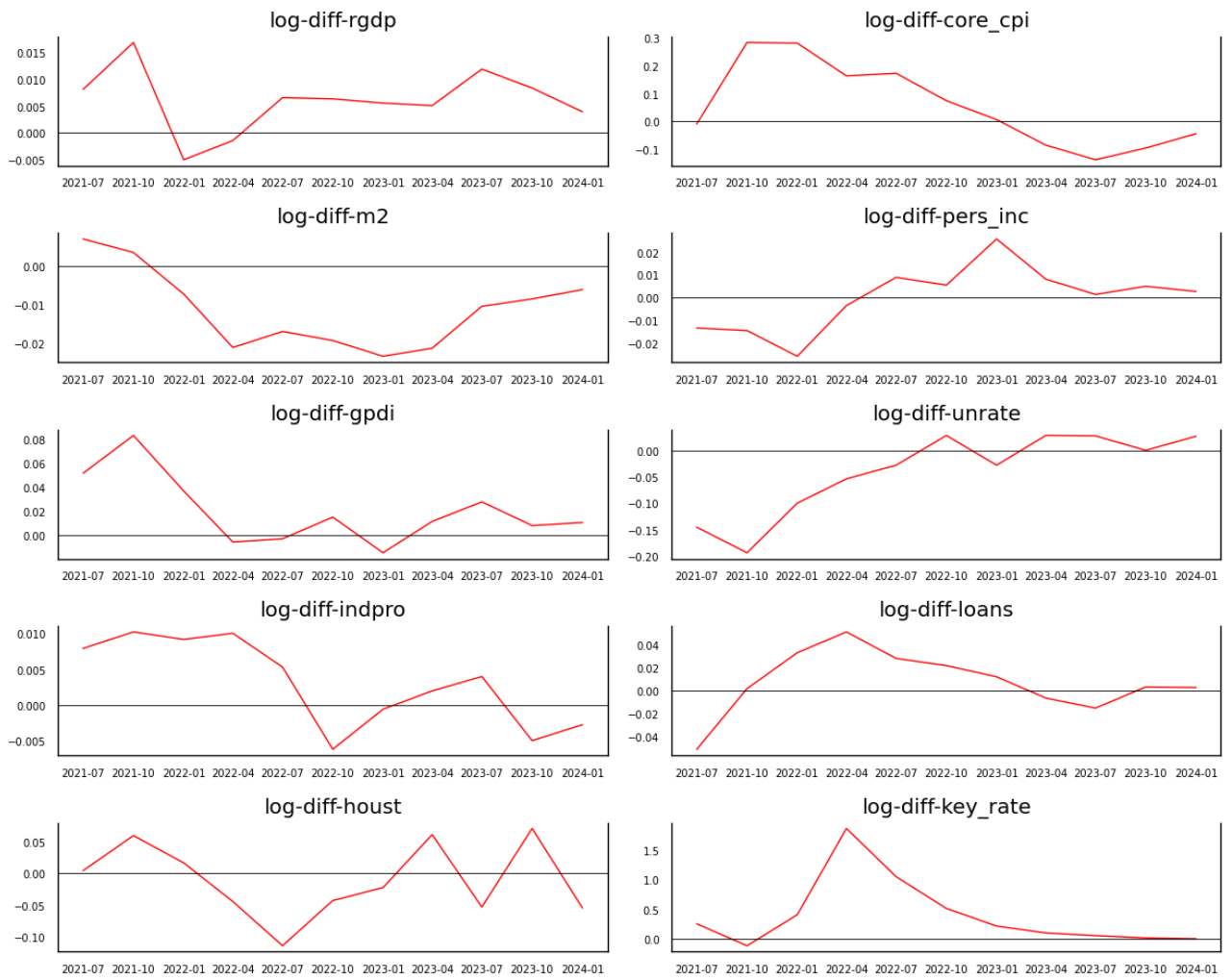


Рисунок 20. Графики преобразованных данных постковидного периода

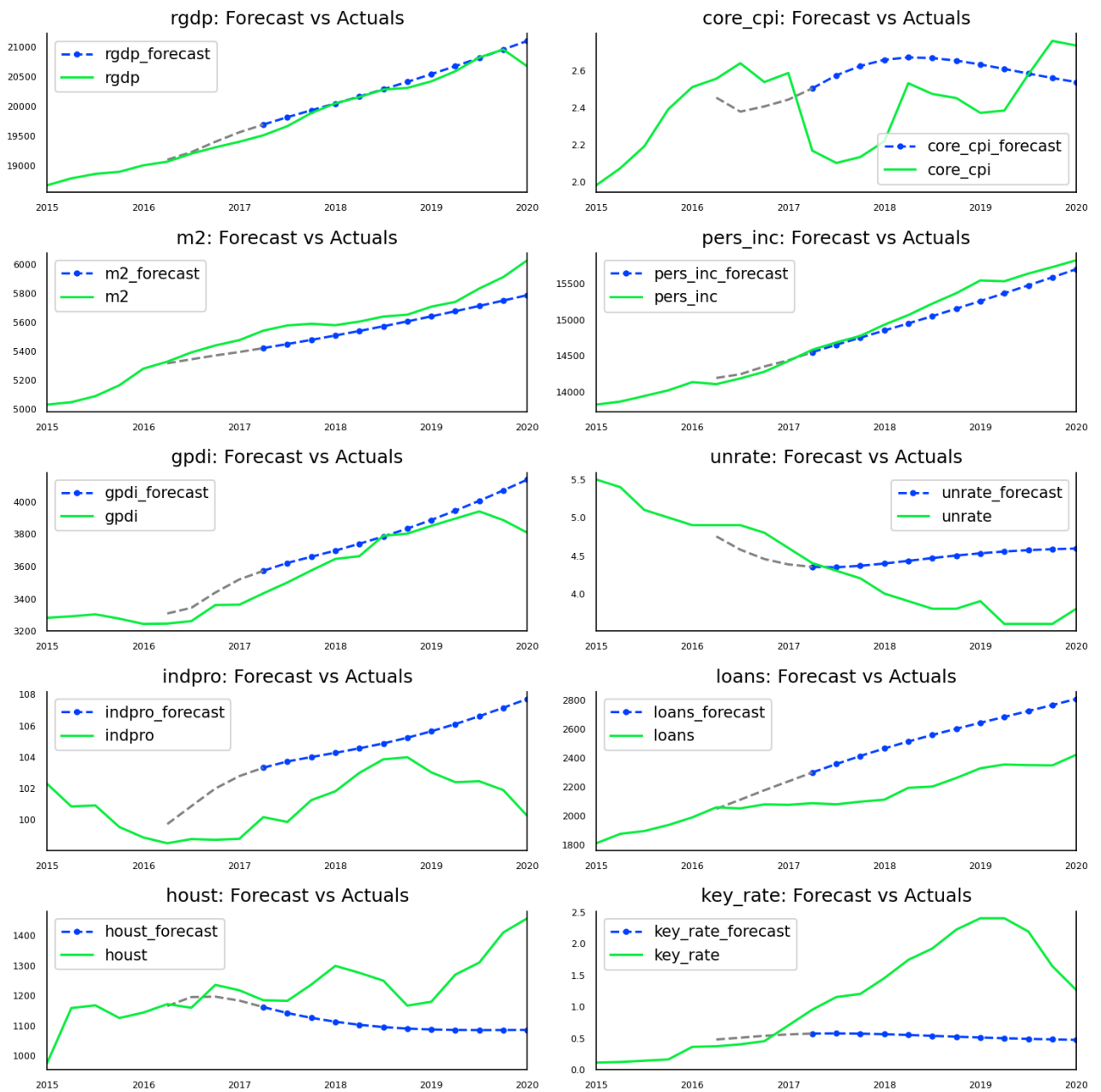


Рисунок 21. Прогнозы модели VAR на доковидном периоде на 16 кварталов

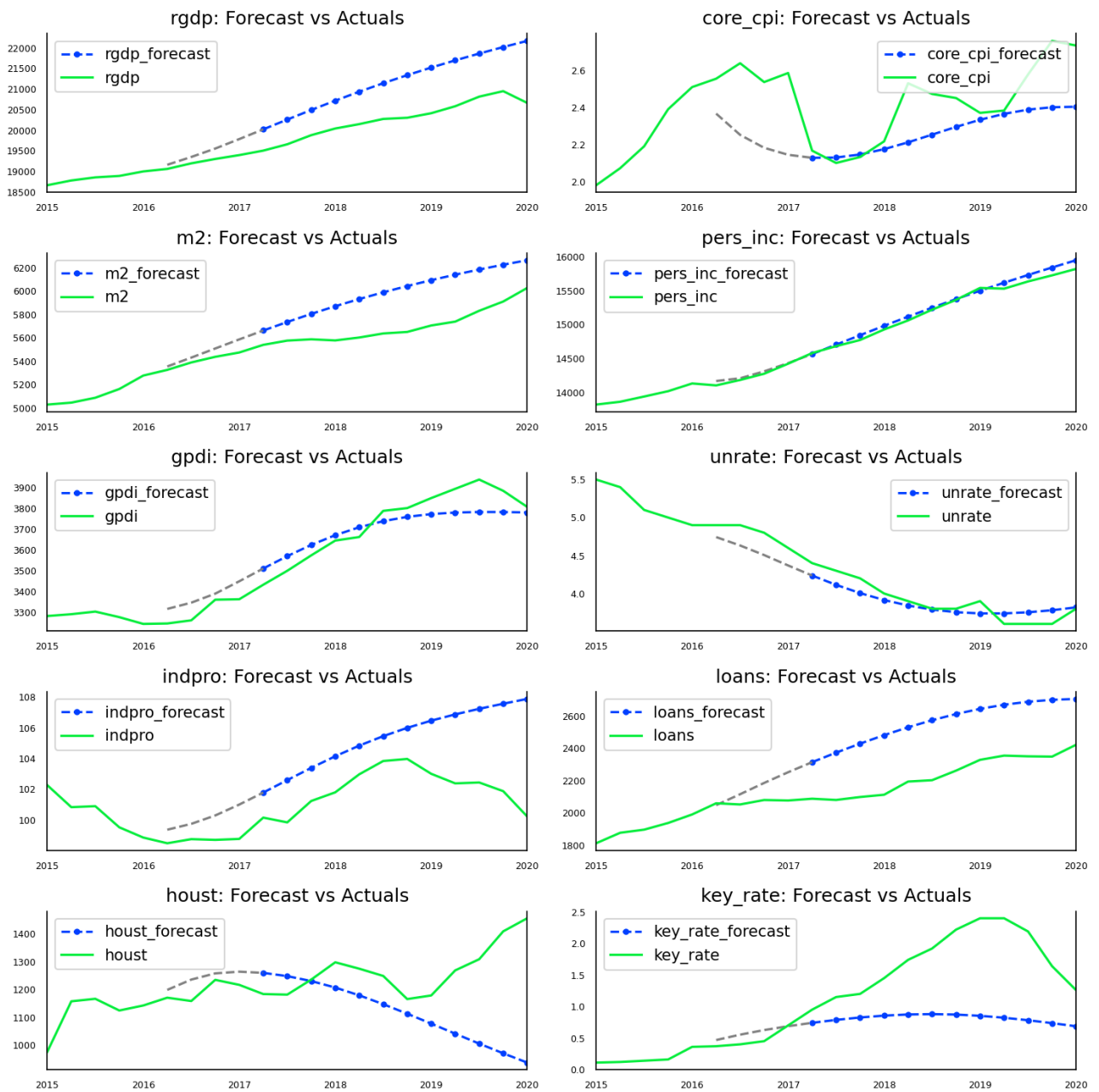


Рисунок 22. Прогнозы модели VECM на доковидном периоде на 16 кварталов

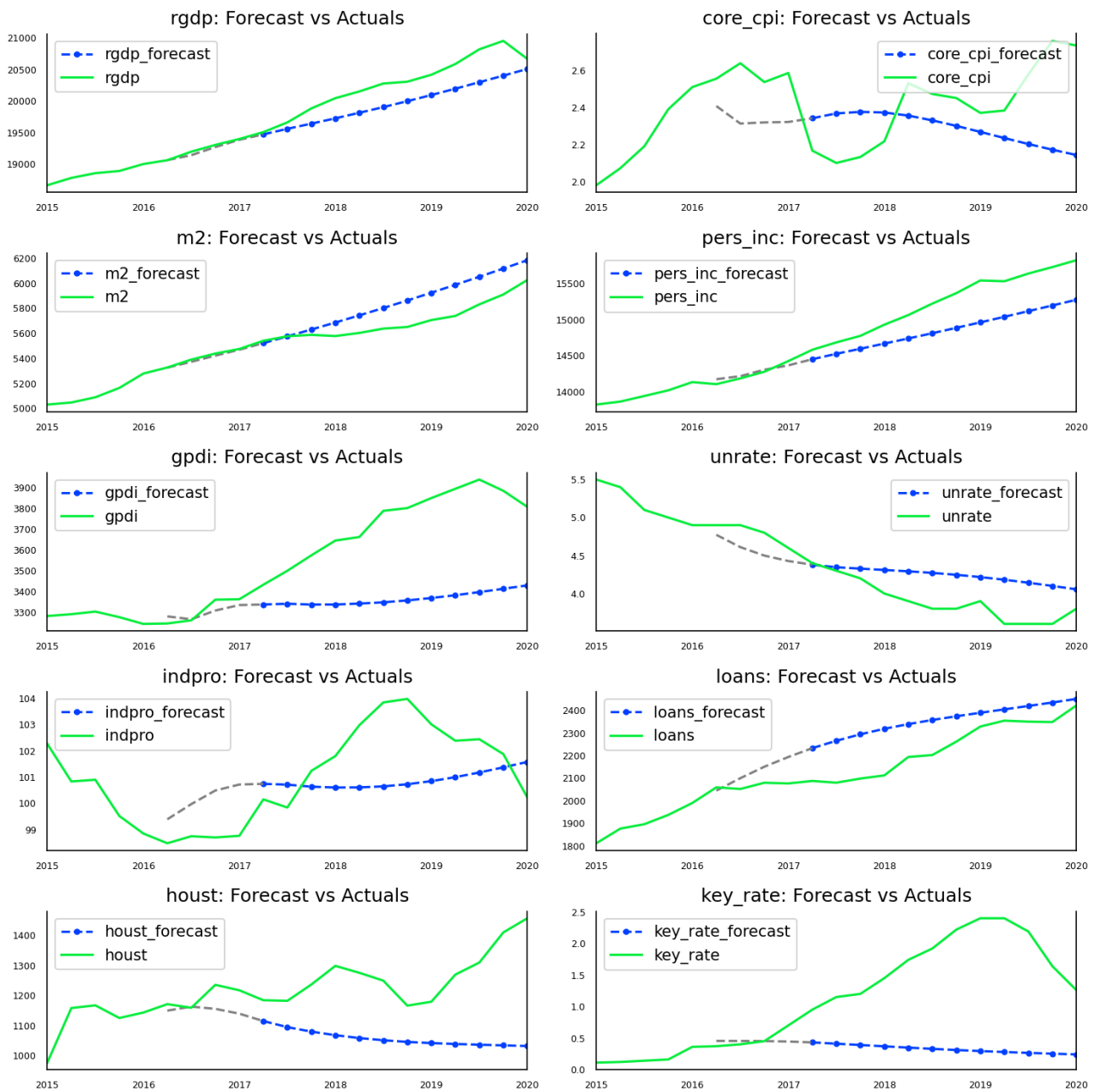


Рисунок 23. Прогнозы модели VARMA на доковидном периоде на 16 кварталов

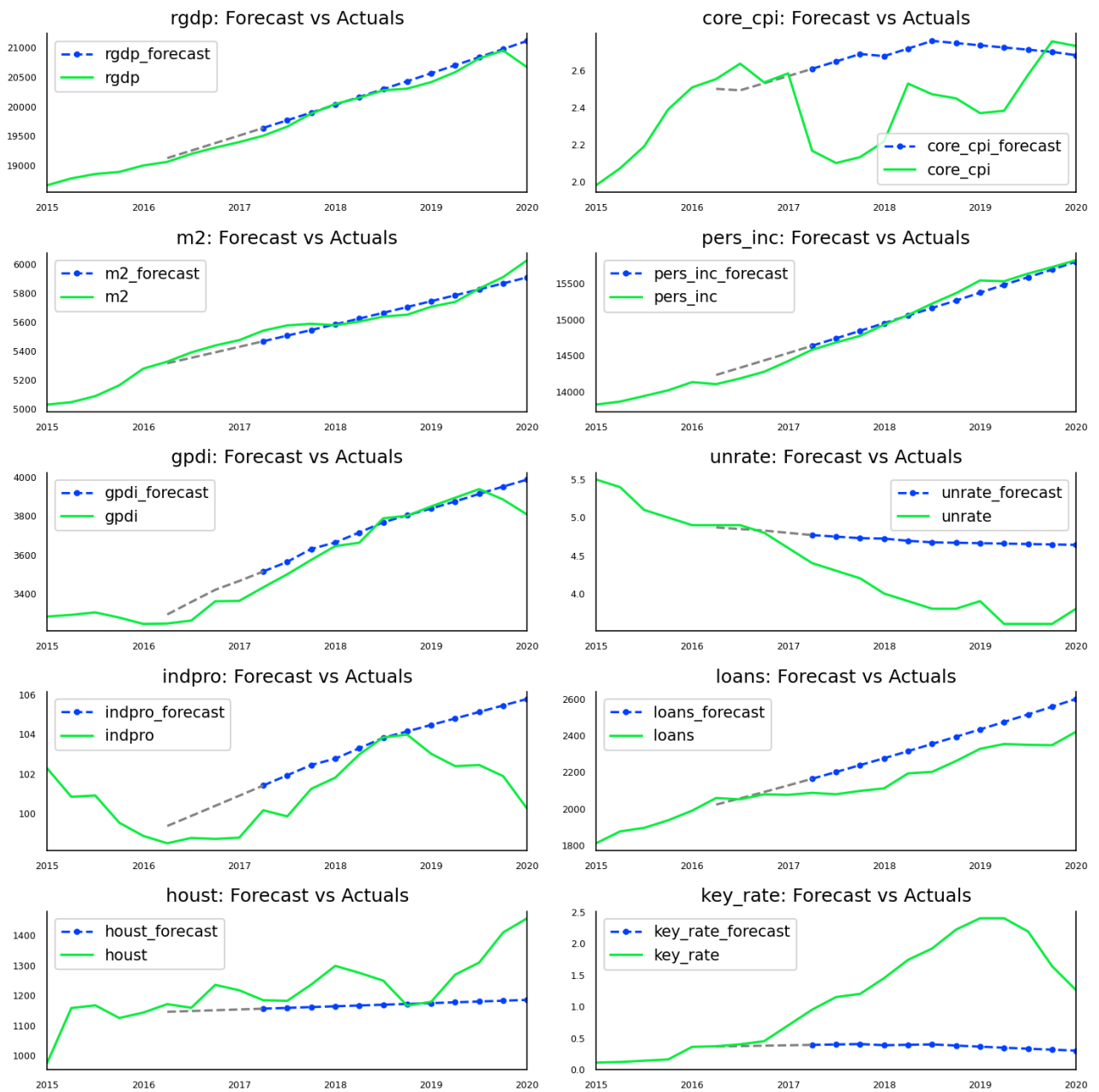


Рисунок 24. Прогнозы модели LightGBM на доковидном периоде на 16 кварталов

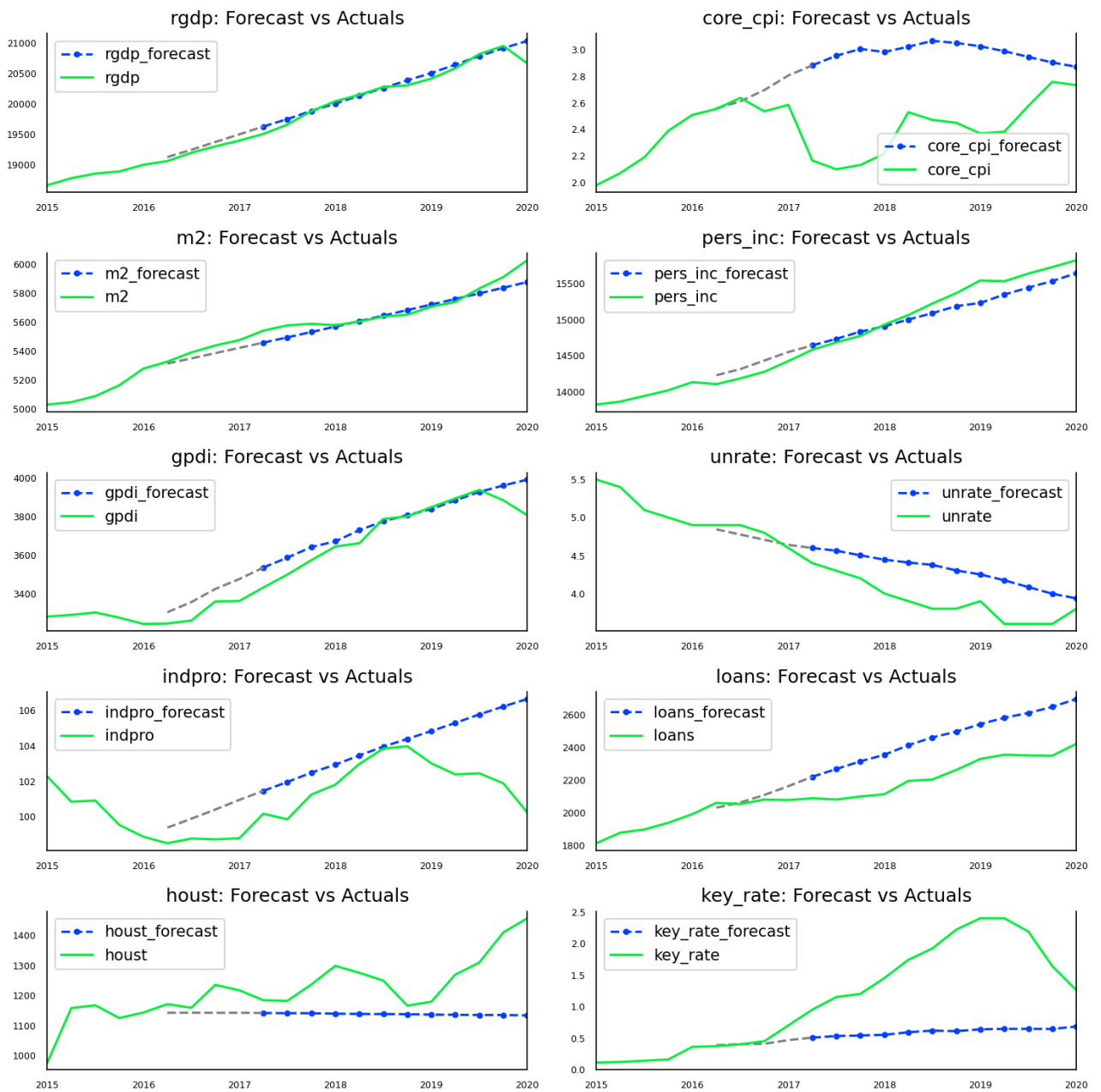


Рисунок 25. Прогнозы модели CatBoost на доковидном периоде на 16 кварталов

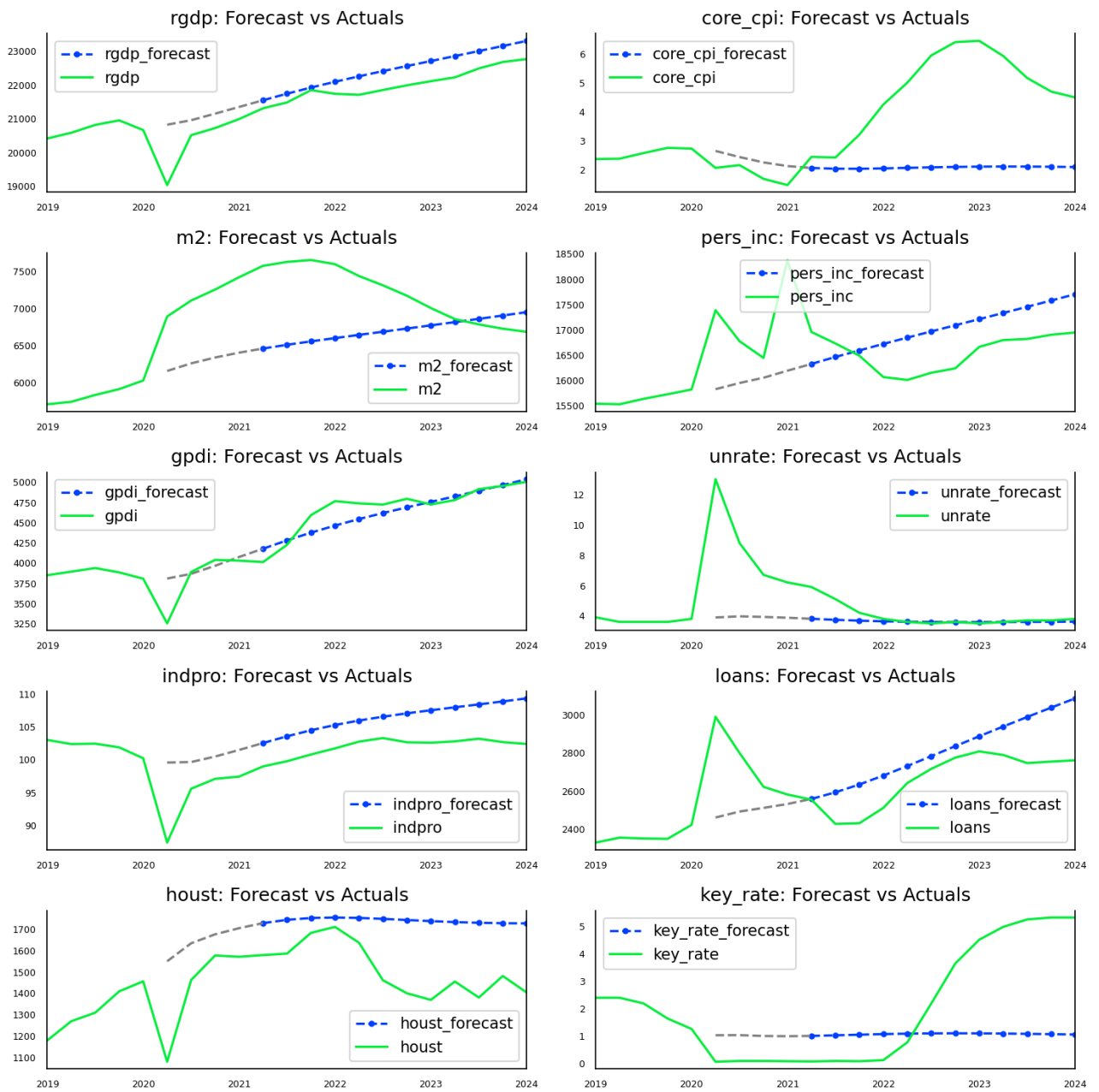


Рисунок 26. Прогнозы модели VAR на постковидный период на 16 кварталов

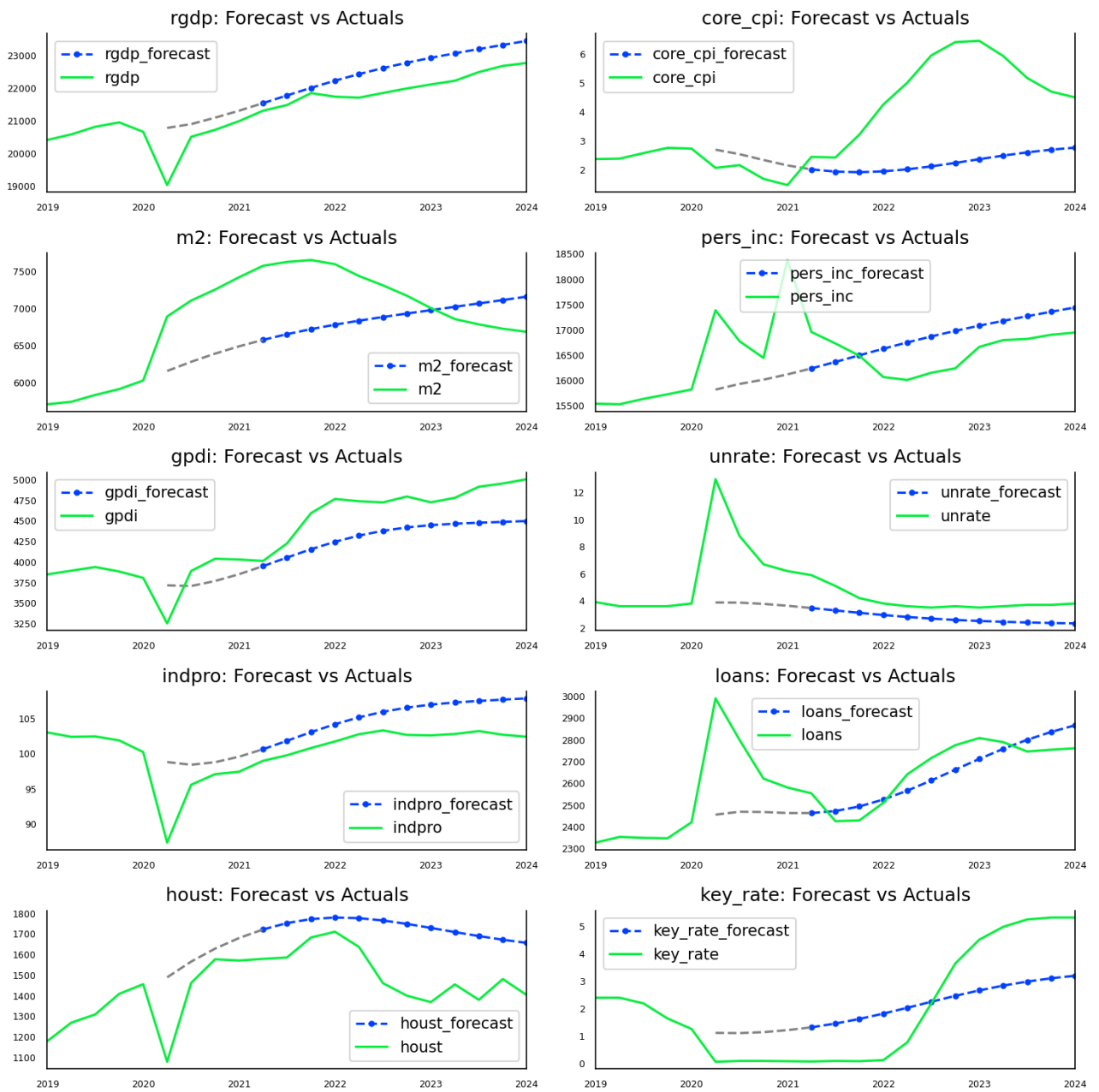


Рисунок 27. Прогнозы модели VECM на постковидный период на 16 кварталов

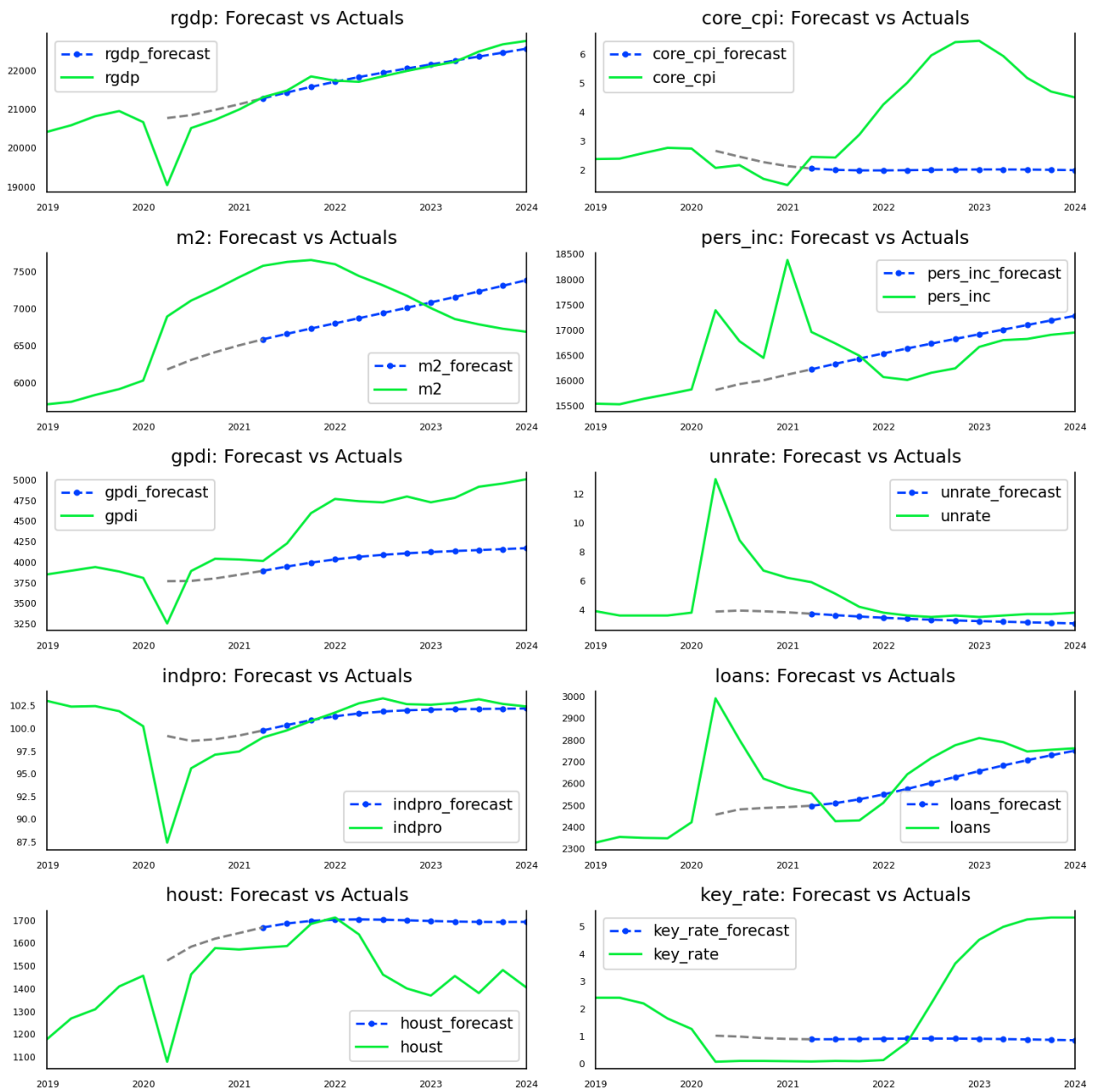


Рисунок 28. Прогнозы модели VARMA на постковидный период на 16 кварталов

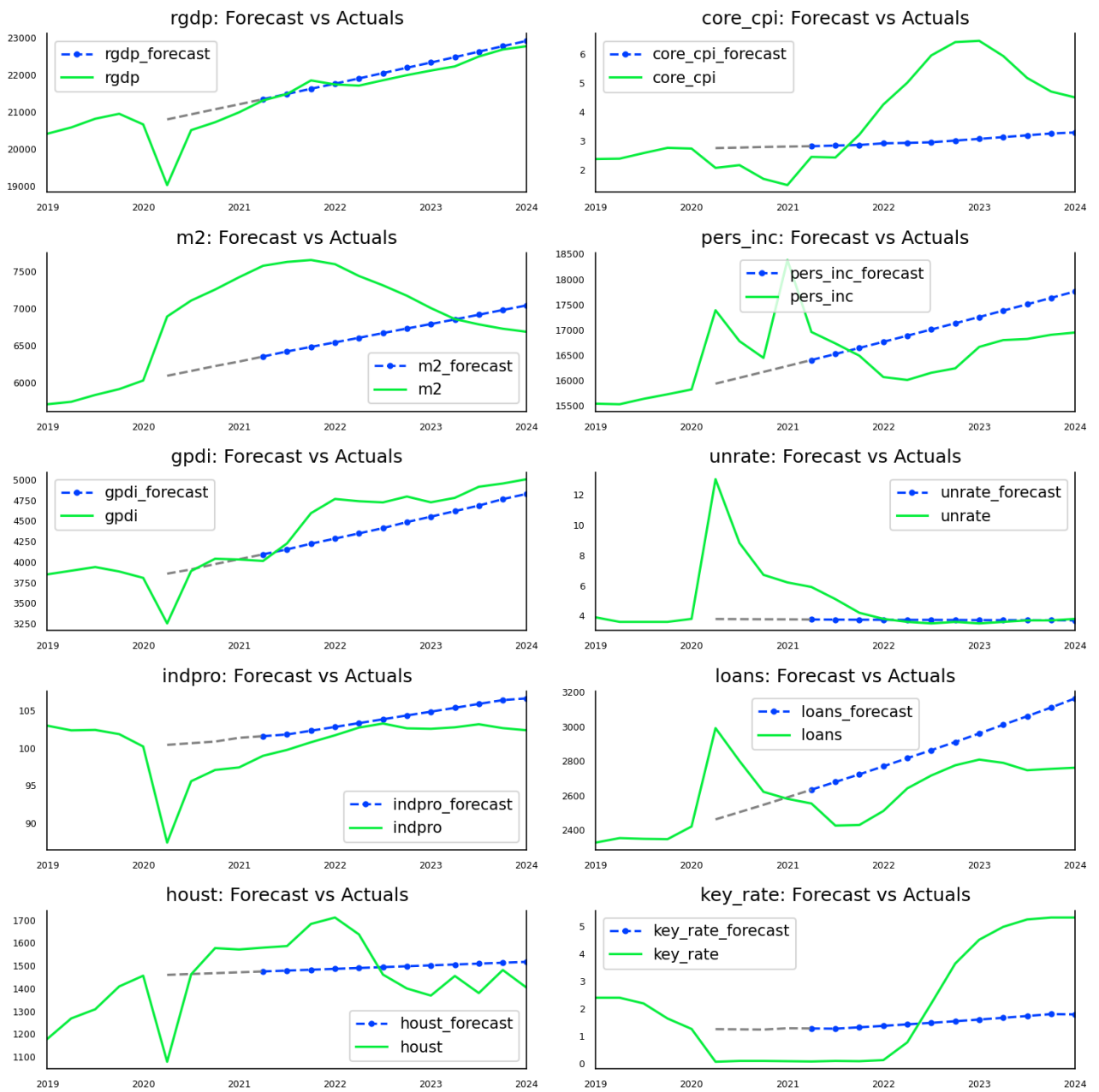


Рисунок 29. Прогнозы модели LightGBM на постковидный период на 16 кварталов

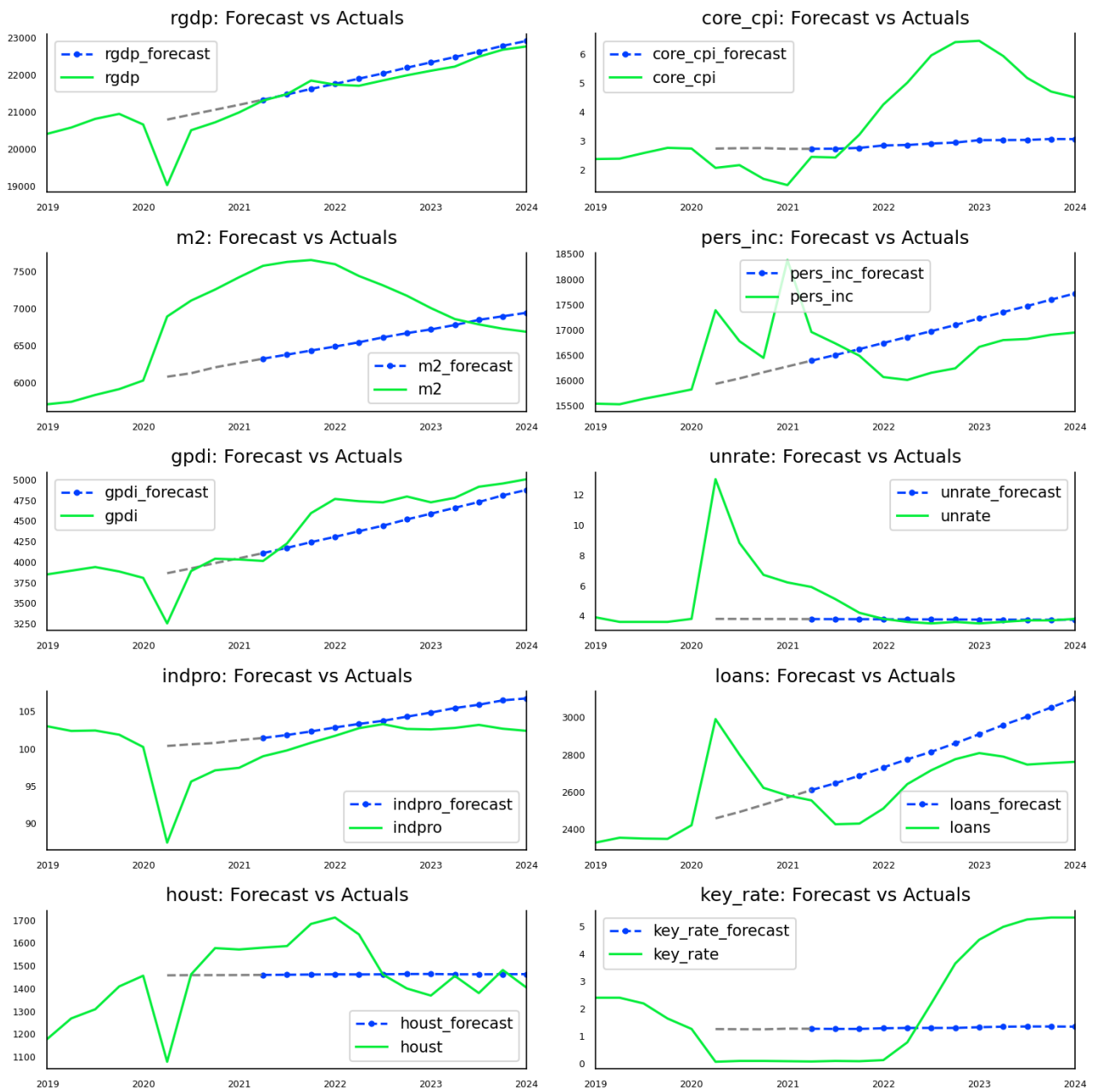


Рисунок 30. Прогнозы модели CatBoost на постковидный период на 16 кварталов

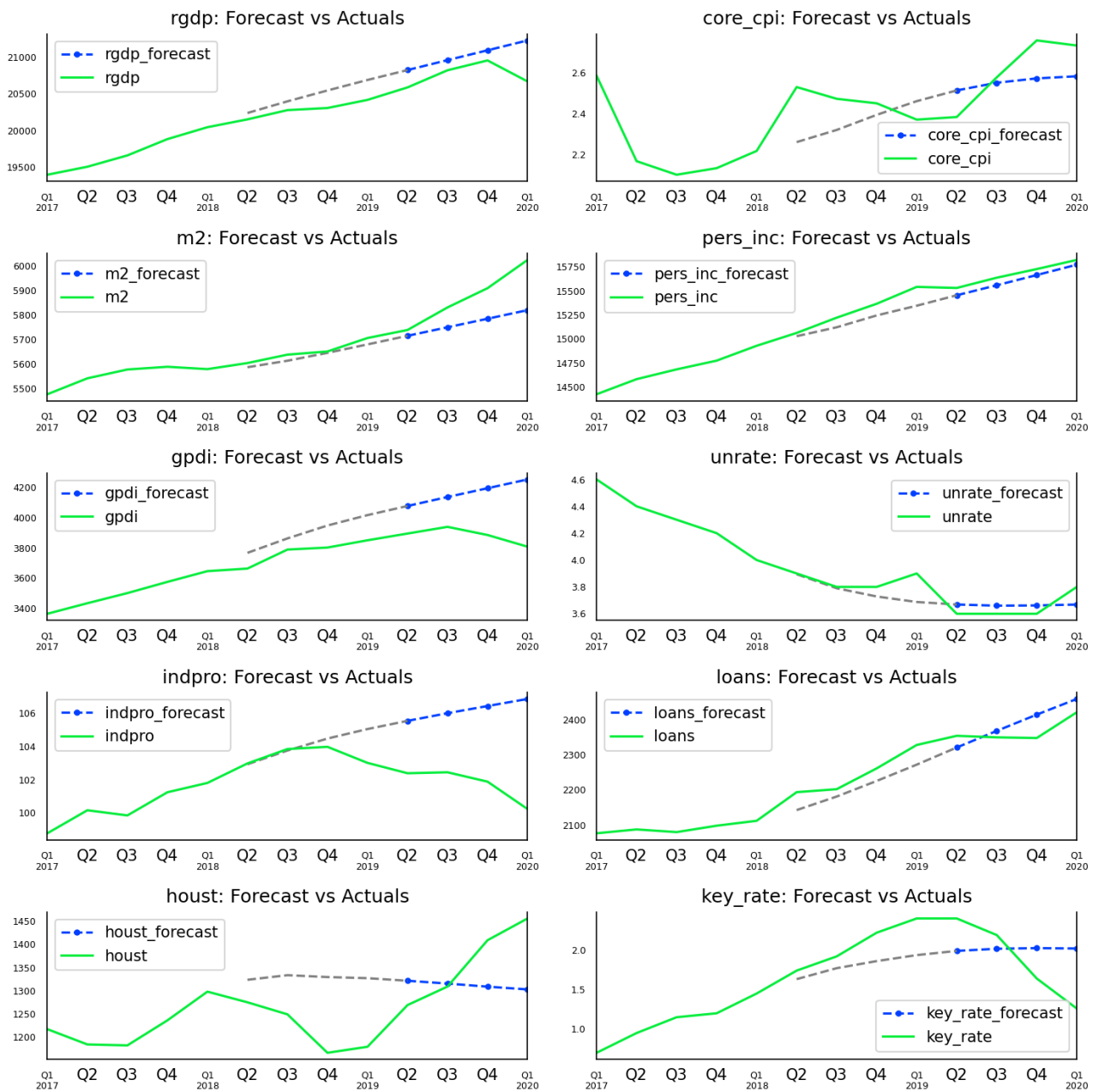


Рисунок 31. Прогнозы модели VAR на доковидный период на 8 кварталов

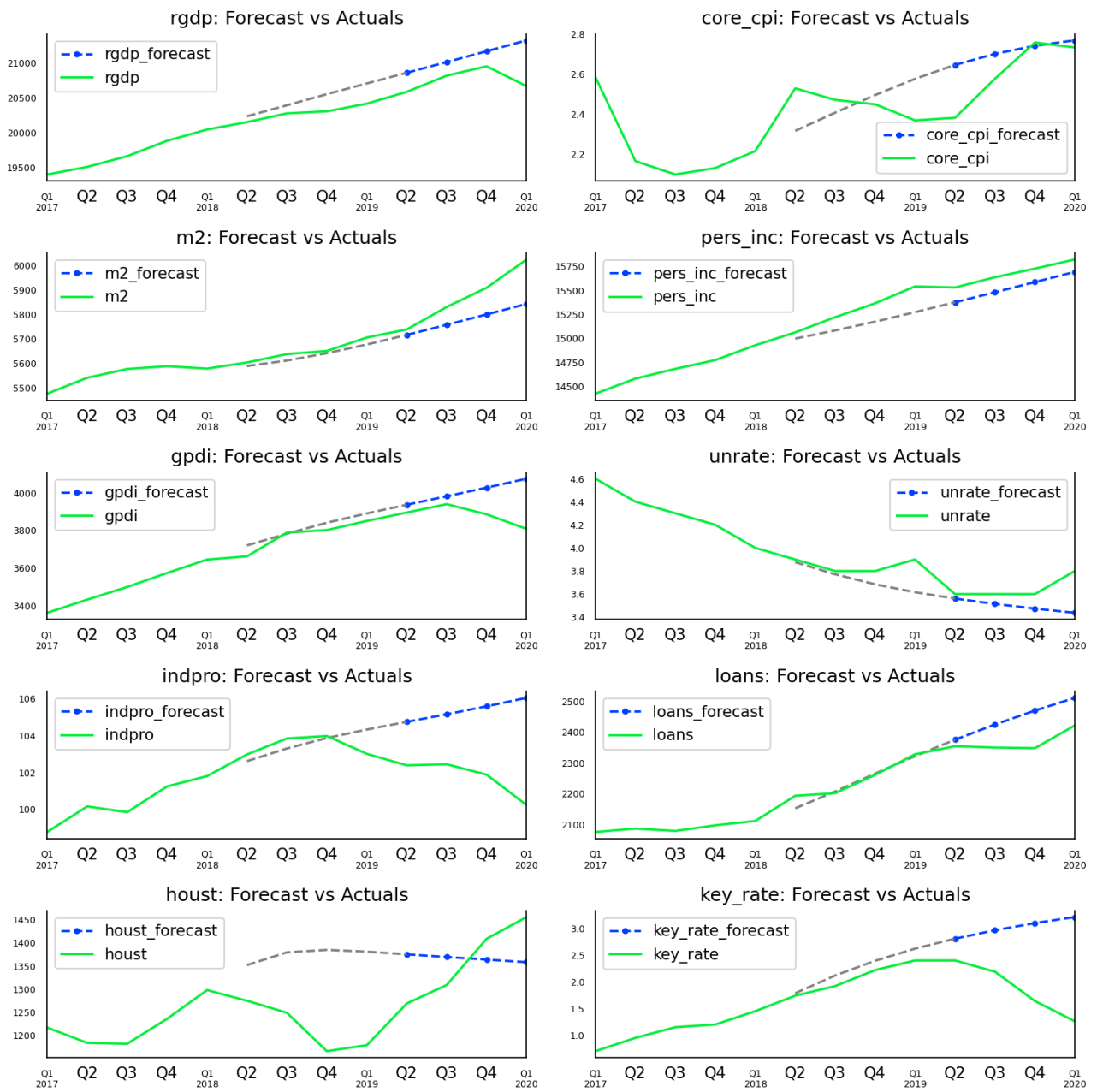


Рисунок 32. Прогнозы модели VECM на доковидный период на 8 кварталов

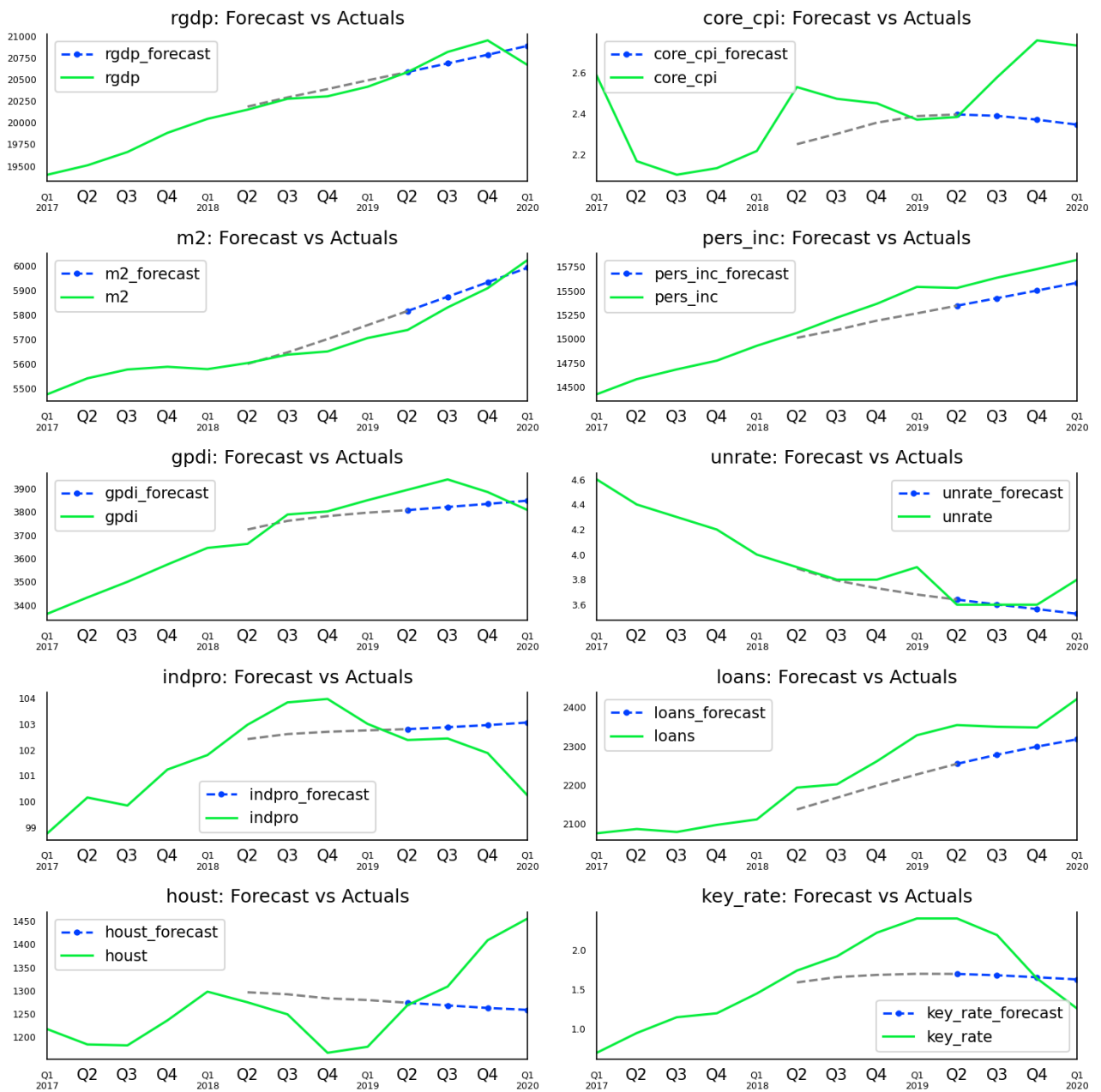


Рисунок 33. Прогнозы модели VARMA на доковидный период на 8 кварталов

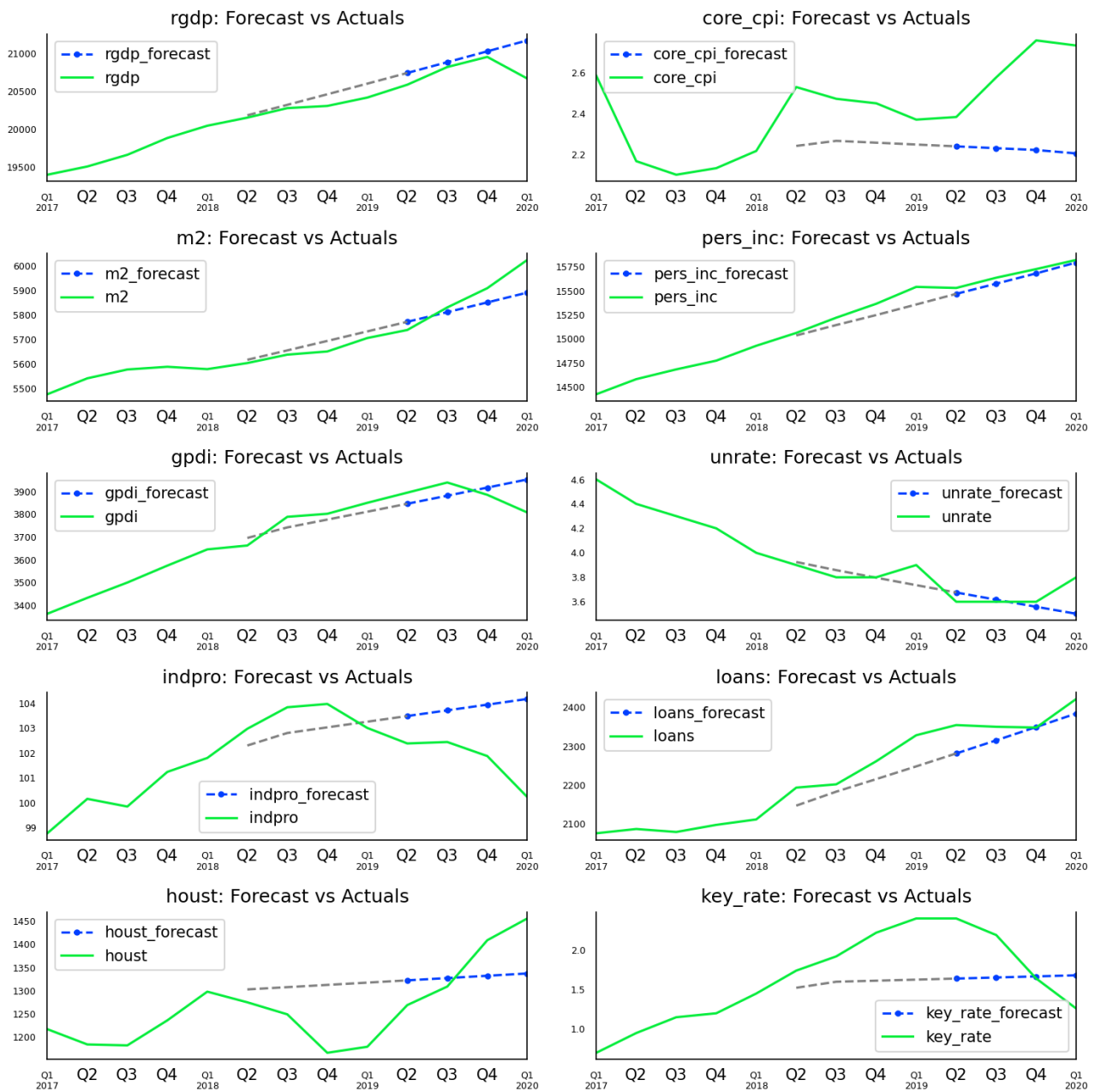


Рисунок 34. Прогнозы модели LightGBM на доковидный период на 8 кварталов

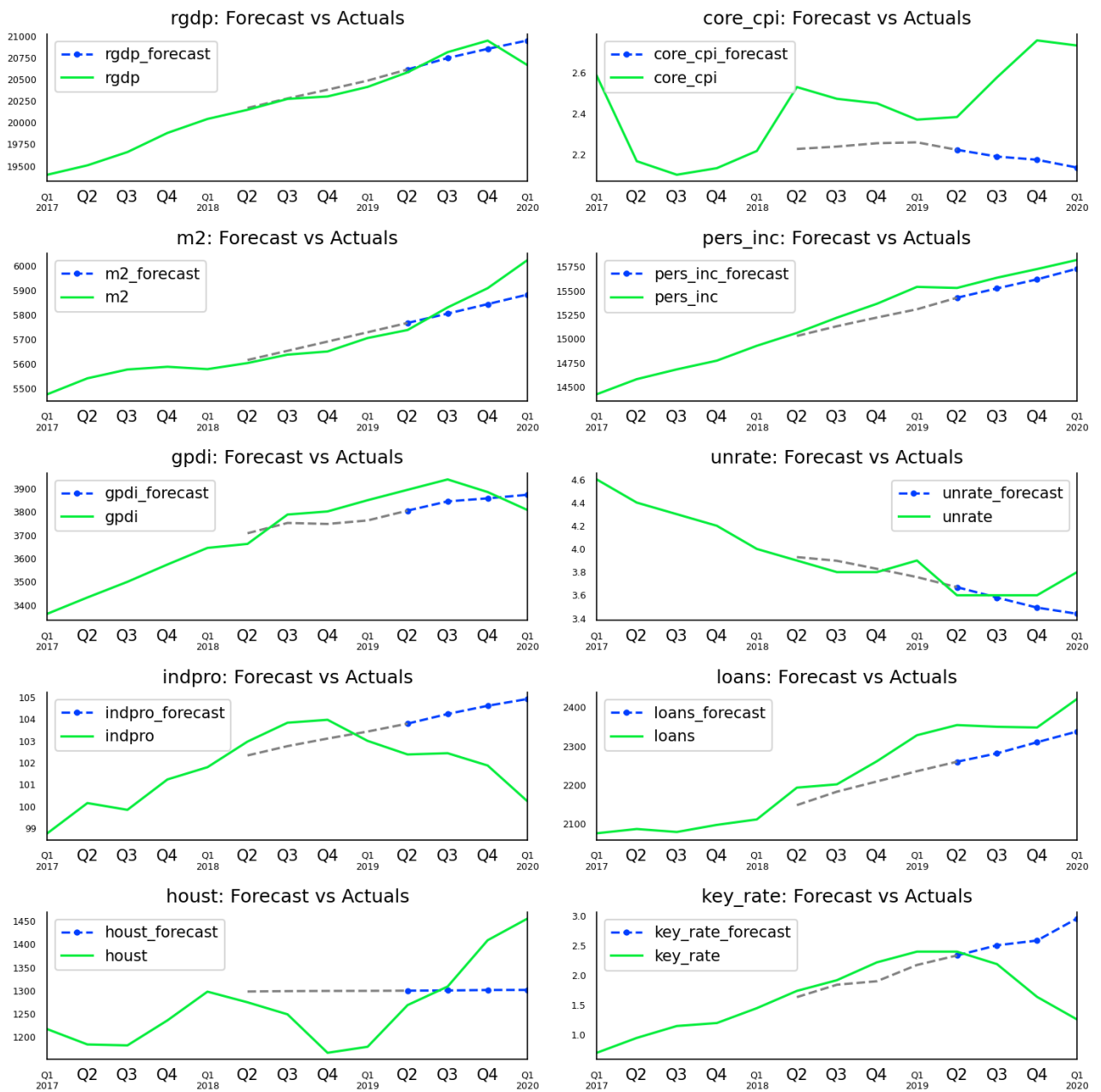


Рисунок 35. Прогнозы модели CatBoost на доковидный период на 8 кварталов

Список пар имеющих разные взаимосвязи в двух периодах, согласно тесту Гренджера. Здесь первая переменная – «причина», вторая – «следствие». Первое число – р-значение в доковидный период, второе – в постковидный.

1. rgdp – core_cpi: 0.0 → 0.3882
2. rgdp – pers_inc: 0.0001 → 0.0176
3. rgdp – gpdi: 0.0 → 0.4074
4. rgdp – loans: 0.0 → 0.9221
5. rgdp – key_rate: 0.0001 → 0.0087
6. core_cpi – rgdp: 0.0001 → 0.0017
7. core_cpi – m2: 0.9538 → 0.0
8. core_cpi – gpdi: 0.0003 → 0.1533
9. core_cpi – unrate: 0.0001 → 0.1117
10. core_cpi – indpro: 0.0001 → 0.5254
11. core_cpi – loans: 0.0065 → 0.0
12. m2 – rgdp: 0.0003 → 0.8105
13. m2 – core_cpi: 0.0011 → 0.0
14. m2 – loans: 0.3294 → 0.0
15. pers_inc – loans: 0.0033 → 0.0
16. gpdi – core_cpi: 0.0 → 0.0185
17. gpdi – pers_inc: 0.0005 → 0.0061
18. unrate – rgdp: 0.0001 → 0.2716
19. unrate – pers_inc: 0.0087 → 0.0004
20. unrate – gpdi: 0.0 → 0.2991
21. unrate – houst: 0.0001 → 0.5159
22. unrate – key_rate: 0.0 → 0.218
23. indpro – rgdp: 0.0 → 0.6437
24. indpro – core_cpi: 0.0 → 0.0515
25. indpro – m2: 0.0003 → 0.6023
26. indpro – pers_inc: 0.0001 → 0.1081
27. indpro – gpdi: 0.0 → 0.6452

- 28.indpro – loans: 0.0 → 0.0045
- 29.indpro – houst: 0.0001 → 0.8915
- 30.loans – gpdi: 0.0059 → 0.0002
- 31.loans – indpro: 0.0004 → 0.5059
- 32.loans – houst: 0.0006 → 0.0002
- 33.houst – rgdp: 0.0 → 0.5174
- 34.houst – gpdi: 0.0 → 0.7995
- 35.houst – unrate: 0.0 → 0.894
- 36.houst – indpro: 0.0 → 0.0043
- 37.key_rate – core_cpi: 0.0 → 0.9356
- 38.key_rate – loans: 0.0 → 0.9198
- 39.key_rate – houst: 0.0192 → 0.0