

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра математического моделирования и анализа данных

БОВТ Тимофей Анатольевич

**КРАТКОСРОЧНОЕ ПРОГНОЗИРОВАНИЕ И НАУКАСТИНГ
МАКРОЭКОНОМИЧЕСКИХ ВРЕМЕННЫХ РЯДОВ НА
ОСНОВЕ ВЕКТОРНЫХ МОДЕЛЕЙ АВТОРЕГРЕССИИ ПО
СМЕШАННЫМ ДАННЫМ**

Дипломная работа

Научный руководитель
док. эконом. наук,
канд. физ.-мат. наук,
профессор В.И. Малюгин

Допущена к защите

«_____» _____ 2025 г.

Зав. кафедрой математического моделирования и анализа данных
док. эконом. наук, канд. физ.-мат. наук,
профессор В.И. Малюгин

Минск, 2025

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет прикладной математики и информатики
Кафедра математического моделирования и анализа данных

УТВЕРЖДАЮ
Заведующий кафедрой

(подпись)(фамилия, инициалы)

14.11.2024 г.

ЗАДАНИЕ
на дипломную работу

Обучающемуся Бовту Тимофею Анатольевичу
Курс 4 Учебная группа 7
Специальность «Прикладная математика»

Тема дипломной работы: «Краткосрочное прогнозирование и наукастинг макроэкономических временных рядов на основе векторных моделей авторегрессии по смешанным данным».

Утверждена указом ректора БГУ от 12.11.2024 № 1295-ПС.

Исходные данные к дипломной работе:

1. Малюгин В. И. Краткосрочное прогнозирование и наукастинг темпов роста инфляции на основе моделей по смешанным данным — Журнал «Банковский вестник» №1/726 — С. 23-36.
2. Макеева, Н.М. Наукастинг элементов использования ВВП России / Н.М. Макеева, И.П. Станкевич – Статья 2022/10, Экономический журнал ВШЭ.
3. Макеева, Н.М. Наукастинг макроэкономических показателей экономики России на основе анализа новостного фона и регулярных данных Росстата / Н. М. Макеева, И. П. Станкевич

Перечень подлежащих разработке вопросов или краткое содержание расчетно-пояснительной записки:

- 1) подготовить обзор основных подходов и методов, применяемых для краткосрочного прогнозирования и наукастинга;
- 2) подготовить описание эконометрических моделей и методов, применяемых для прогнозирования прироста ВВП Республики Беларусь;
- 3) разработать алгоритмический инструментарий и программное обеспечение

на основе метода расширяющегося окна для краткосрочного прогнозирования по моделям векторной авторегрессии по смешанным данным;

4) провести сравнительный анализ моделей по смешанным данным с моделями по агрегированным данным;

Примерный календарный график выполнения дипломной работы:

- **февраль (1-ая неделя)** – получение задания, загрузка новых данных для исследования; изучение функции импульсных откликов для VAR модели;
- **февраль (2-3-я неделя)** – сбор статистик по ВВП, опережающим показателям; изучение основных статистических методов работы с экономическими данными;
- **март (4-5-ая неделя)** – изучение основных методов расчета ВВП, изучение экономических взаимосвязей между ВВП и опережающими показателями; исследование показателя СИЭН, его взаимосвязи с ВВП; построение плана оформления I-II глав дипломной работы;
- **март (6-7-ая неделя)** – написание отчета по теоретическим взаимосвязям модели ВВП и статистическому анализу временных рядов; программирование алгоритма расширяющегося окна; оценка точности прогнозов;
- **апрель (8-9-ая неделя)** – составление плана III главы в дипломной работе; общая характеристика модели ВВП белорусской экономики; написание III главы, сравнительный анализ построенных моделей;
- **апрель (10-11-ая неделя)** – оформление отчета по проделанной работе; интерпретация полученных результатов;
- **май (12-15-ая неделя)** – подготовка презентации, прохождение системы «Антиплагиат», получение допуска, рецензирование дипломной работы.

Дата выдачи задания 14.11.2024.

Срок сдачи законченной дипломной работы 27.05.2025.

Руководитель дипломной работы _____ В. И. Малюгин
(подпись)

Подпись обучающегося _____
(подпись)

Дата 14.11.2024 г.

Проинформирован о недопустимости привлечения третьих лиц к выполнению дипломной работы, плагиата, фальсификации или подлога материалов.

_____ Т. А. Бовт
(подпись)

ОГЛАВЛЕНИЕ

РЕФЕРАТ	3
ВВЕДЕНИЕ.	6
ГЛАВА 1 ОБЗОР ЭКОНОМИКО-МАТЕМАТИЧЕСКИХ МЕТОДОВ И ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ ВЫЧИСЛЕНИЯ И ПРОГНОЗИРОВАНИЯ РЕАЛЬНОГО ВВП	8
1.1 Введение в проблематику макроэкономического моделирования ВВП.	8
1.2 Методы расчета валового внутреннего продукта и сроки выхода оценок.	9
1.3 Доли вклада компонент в ВВП Беларуси	13
1.4 Опыт краткосрочного прогнозирования и наукастинга реального ВВП на основе моделей по данным смешанной частоты	17
ГЛАВА 2 МНОГОМЕРНЫЕ МОДЕЛИ И МЕТОДЫ АНАЛИЗА ПО ДАННЫМ СМЕШАННОЙ ЧАСТОТЫ	22
2.1 Модель векторной авторегрессии VAR и ее применение в задачах макроэкономического анализа	22
2.2 Модель векторной авторегрессии по данным смешанной частоты MF-VAR.	28
2.3 Оценка точности модельных прогнозов	30
ГЛАВА 3 ПРОГНОЗИРОВАНИЕ ТЕМПОВ ПРИРОСТА РЕАЛЬНОГО ВВП НА ОСНОВЕ МОДЕЛИ MF-VAR	33
3.1 Постановка задачи и описание исследуемых макроэкономических показателей	33
3.2 Предварительный анализ временных рядов	40
3.3 Теоретическое описание макроэкономической модели ВВП	44
3.4 Результаты построения и использования моделей в задачах экономического анализа и прогнозирования	47
ЗАКЛЮЧЕНИЕ	58
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	59

РЕФЕРАТ

Дипломная работа включает 62 страницы, 33 рисунка, 9 таблиц, 43 источника.

Ключевые слова: ДАННЫЕ СМЕШАННОЙ ЧАСТОТЫ; КРАТКОСРОЧНОЕ ПРОГНОЗИРОВАНИЕ И НАУКАСТИНГ; МОДЕЛЬ MF-VAR; ПРОГНОЗИРОВАНИЕ ПРИРОСТА РЕАЛЬНОГО ВВП; МАКРОЭКОНОМИЧЕСКИЕ ПОКАЗАТЕЛИ; ОПЕРЕЖАЮЩИЕ ПОКАЗАТЕЛИ; БЕЛОРУССКАЯ ЭКОНОМИКА.

Объект исследования – многомерные эконометрические модели по смешанным данным и их применение в задачах прогнозирования макроэкономических показателей.

Цель исследования – разработка модельного и алгоритмического инструментария для краткосрочного прогнозирования и наукастинга; оценка точности прогнозов на основе многомерных эконометрических моделей по смешанным данным.

Методы исследования – методы теории вероятностей и математической статистики, методы многомерных временных рядов, методы эконометрического анализа, методы макроэкономического моделирования.

Полученные результаты и их новизна – подготовлен обзор основных методов краткосрочного прогнозирования и наукастинга; подготовлено описание эконометрических моделей и методов для прогнозирования прироста ВВП Республики Беларусь; разработан алгоритмический инструментарий на основе метода расширяющегося окна для краткосрочного прогнозирования по MF-VAR моделям; проведен сравнительный анализ моделей по смешанным данным с моделями по агрегированным данным. Новизна заключается в том, что впервые были построены MF-VAR модели на белорусской экономике.

Достоверность материалов и результатов дипломной работы. Результаты исследований получены на основе данных Национального статистического комитета Республики Беларусь, согласованы со специалистами Национального Банка Республики Беларусь, а также согласуются с остальными исследованиями по белорусской экономике.

Область возможного практического применения – краткосрочное прогнозирование и наукастинг реального ВВП Беларуси для последующего сравнения с другими моделями.

РЭФЭРАТ

Дыпломная праца ўключае 62 старонкі, 33 малюнкаў, 9 табліц, 43 крыніцы.

Ключавыя словы: ДАДЗЕННЫЯ ЗМЯШАНАЙ ЧАСТОТЫ; КАРОТКАТЭРМІНОВАЕ ПРАГНАЗАВАННЕ І НАУКАСТІНГ; МАДЭЛЬ MF-VAR; ПРАГНАЗАВАННЕ ПРЫРОСТУ РЭАЛЬНАГА ВУП; МАКРАЭКАНАМІЧНЫЯ ПАКАЗЧЫКІ; АПЕРАДЖАЛЬНЫЯ ПАКАЗЧЫКІ; БЕЛАРУСКАЯ ЭКАНОМІКА.

Аб'ект даследавання – шматмерныя эканаметрычнага мадэлі па змяшаным дадзеных і іх прымяненне ў задачах прагназавання макраэканамічных паказчыкаў.

Цэль даследавання – распрацоўка мадэльнага і алгарытмічнага інструментара для кароткатэрміновага прагназавання і наукастынга; ацэнка дакладнасці прагнозаў на аснове шматмерных эканаметрычнага мадэляў па змяшаным дадзеных.

Метады даследавання – метады тэорыі верагоднасцяў і матэматычнай статыстыкі, метады шматмерных часовых шэрагаў, метады эканаметрычнага аналізу, метады макраэканамічнага мадэлявання.

Атрыманыя вынікі і іх навізна – падрыхтаваны агляд асноўных метадаў кароткатэрміновага прагназавання і наукастынга; падрыхтавана Апісанне эканаметрычнага мадэляў і метадаў для прагназавання прыросту ВУП Рэспублікі Беларусь; распрацаваны алгарытмічны інструментар на аснове метаду пашыраецца вокны для кароткатэрміновага прагназавання па MF-VAR мадэлям; праведзены параўнальны аналіз мадэляў па змяшаным дадзеных з мадэлямі па агрегаваных дадзеных. Упершыню былі пабудаваны мадэлі MF-VAR на беларускай эканоміцы.

Дакладнасць матэрыялаў і вынікаў дыпломнай працы. Вынікі даследаванняў атрыманы на аснове даных Нацыянальнага статыстычнага камітэта Рэспублікі Беларусь, узгоднены са спецыялістамі Нацыянальнага Банка Рэспублікі Беларусь, а таксама ўзгадняюцца з астатнімі даследаваннямі па беларускай эканоміцы.

Вобласць магчымага практычнага прымянення – кароткатэрміновае прагназаванне і наукастынг рэальнага ВУП Беларусі для наступнага параўнання з іншымі мадэлямі.

ABSTRACT

Graduate work includes 62 pages, 33 figures, 9 tables, 43 references.

Key words: MIXED FREQUENCY DATA; SHORT-TERM FORECASTING AND SCIENCASTING; MF-VAR MODEL; FORECASTING REAL GDP GROWTH; MACROECONOMIC INDICATORS; LEADING INDICATORS; BELARUSIAN ECONOMY.

Object of research – multidimensional econometric models based on mixed data and their application in forecasting macroeconomic indicators.

Purpose of research – development of model and algorithmic tools for short-term forecasting and sciencasting; estimation of forecast accuracy based on multidimensional econometric models based on mixed data.

Research methods – methods of probability theory and mathematical statistics, methods of multidimensional time series, methods of econometric analysis, methods of macroeconomic modeling.

Obtained results and their novelty. An overview of the main methods of short-term forecasting and sciencasting has been prepared; a description of econometric models and methods for predicting GDP growth in the Republic of Belarus has been prepared; algorithmic tools based on the expanding window method for short-term forecasting using MF-VAR models have been developed; a comparative analysis of models based on mixed data with models based on aggregated data has been carried out. The novelty lies in the fact that MF-VAR models were built on the Belarusian economy for the first time.

Authenticity of the materials and results of the diploma work. The research results were obtained on the basis of data from the National Statistical Committee of the Republic of Belarus, coordinated with specialists from the National Bank of the Republic of Belarus, and are also consistent with other studies on the Belarusian economy.

Area of possible practical application – short-term forecasting and sciencasting of the real GDP of Belarus for subsequent comparison with other models.

ВВЕДЕНИЕ

Цель данной дипломной работы – разработка модельного и алгоритмического инструментария для краткосрочного прогнозирования и наукастинга, а также оценка точности прогнозов на основе многомерных эконометрических моделей по смешанным данным.

Отдельные фундаментальные макроэкономические показатели формируются Национальным статистическим комитетом Республики Беларусь на квартальной и годовой частоте. В то же время доступны более оперативные показатели, которые публикуются с месячной и дневной периодичностью. В частности, показатель валового внутреннего продукта (ВВП) по методу использования доходов формируется на квартальной частоте, а отраслевые показатели и статистика индексов цен – на месячной частоте, обменные курсы валют и денежно-кредитные показатели – на дневной частоте. Причем в соответствии с Регламентом публикации данных оценок ВВП Республики Беларусь квартальная оценка ВВП публикуется на 90-ый день после окончания отчетного квартала. В связи с этим становится актуальным вопрос об использовании более оперативной информации при прогнозировании показателя ВВП [8, 9].

Все классические регрессионные модели машинного обучения работают с данными, заданными на одной частоте. Соответственно в ходе предварительного анализа приходится преобразовывать данные к одной частоте. Как правило, для этого используется агрегация данных высокой частоты. Например, если исследуемая эндогенная переменная находится в квартальном представлении, а экзогенные факторы – в месячном, то мы можем составить новый набор экзогенных переменных, взяв в качестве квартального значения последний месяц, сумму или среднее арифметическое месячных значений за текущий квартал. Таким образом, мы получим все данные в одной частоте, что позволяет нам использовать большое количество моделей машинного обучения для моделирования необходимого нам показателя.

Однако такой подход имеет значительный недостаток: возникает потеря некоторой информации о динамике объясняющих данных, которая может быть крайне полезна при построении модели.

Таким образом, вопрос о том, как можно без преобразования данных и потери какой-либо информации строить регрессионную модель для моделирования исследуемых показателей, становится актуален.

В данной работе для решения задачи моделирования показателя ВВП

Республики Беларусь, представленного на квартальной частоте, по опережающим факторам на месячной частоте рассматриваются специальные регрессионные модели, предназначенные для работы с данными смешанной частоты – модели MF-VAR (Mixed Frequency Vector Autoregression). Они используют особый способ агрегации высокочастотных факторов таким образом, что в модели учитывается каждое известное значение факторов. В последнее время эти модели используются для прогнозирования макроэкономических временных рядов, где обычно квартальный рост ВВП прогнозируется по месячным макроэкономическим и финансовым показателям. Векторные авторегрессионные модели позволяют не только установить характер зависимости эндогенного показателя от экзогенных, но и понять, как в целом построена связь между всеми переменными в системе и их поведением в прошлом. Разработка моделей, способных работать с переменными, отбираемыми на разной частоте вызывает значительный интерес в сфере эконометрии [30].

Для достижения цели решаются следующие задачи:

- 1) подготовка обзора основных подходов и методов, применяемых для краткосрочного прогнозирования и наукастинга;
- 2) подготовка описания эконометрических моделей и методов, применяемых для прогнозирования простора ВВП Республики Беларусь;
- 3) разработка алгоритмического инструментария и программного обеспечения на основе метода расширяющегося окна для краткосрочного прогнозирования по моделям векторной авторегрессии по смешанным данным;
- 4) проведение сравнительного анализа моделей по смешанным данным с моделями по агрегированным данным.

Работа имеет следующую структуру:

- 1) в первой главе подготовлен обзор экономико-математических методов и эконометрических моделей вычисления и прогнозирования реального ВВП;
- 2) во второй главе подготовлено описание многомерных моделей и методов анализа по данным смешанной частоты, применяемых для прогнозирования прироста ВВП Республики Беларусь;
- 3) в третьей главе подготовлено описание результатов прогнозирования темпов прироста реального ВВП на основе модели MF-VAR.

ГЛАВА 1

ОБЗОР ЭКОНОМИКО-МАТЕМАТИЧЕСКИХ МЕТОДОВ И ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ ВЫЧИСЛЕНИЯ И ПРОГНОЗИРОВАНИЯ РЕАЛЬНОГО ВВП

В данной главе формулируются основные понятия, используемые в работе. Приводятся различные методы расчета ВВП, применяемые в экономической науке. Оценивается доля компонент ВВП. Описывается опыт решения задачи наукастинга по данным смешанной частоты.

1.1 Введение в проблематику макроэкономического моделирования ВВП

Макроэкономическое моделирование занимает ключевое место в анализе и прогнозировании национальной экономики, поскольку позволяет выявлять закономерности функционирования экономической системы, оценивать влияние различных факторов на экономический рост и разрабатывать эффективные меры экономической политики [8, 9].

В настоящее время общепринятым ключевым показателем работы национальной экономики выступает *валовой внутренний продукт* (ВВП). Данный макроэкономический показатель отражает общую денежную стоимость всех конечных товаров и услуг, произведенных за определенный период времени во всех отраслях экономики на территории конкретного государства вне зависимости от национальной принадлежности использованных факторов производства [1, 3].

ВВП является наиболее часто используемым показателем общей экономической активности государства. Он применяется как для оценки роста или падения экономического состояния конкретной страны, так и для замеров общемировой экономики и для сравнения экономик стран между собой. Показатели ВВП влияют на многие сферы жизни в стране. По повышению или понижению компонент ВВП мы можем судить о том, насколько все хорошо или плохо в экономике страны.

В экономической науке разделяют номинальный и реальный ВВП.

Номинальный ВВП, или ВВП в текущих ценах, – это стоимость товаров и услуг, рассчитанная по ценам, по которым они фактически производятся и продаются в текущем году, без какой-либо поправки на инфляцию. Вследствие этого номинальный ВВП зависит от изменений индекса цен и доходов рассматриваемой экономики.

Реальный ВВП – это стоимость товаров и услуг, рассчитанная по ценам базового года, что позволяет исключить из номинального ВВП инфляцию/дефляцию и сравнить реальные объемы производства в разные годы. В реальном ВВП учитывается, в какой степени рост ВВП определяется реальным ростом производства, а не ростом цен [19].

Исходя из определения реального ВВП, имеем соотношение

$$Y_t^{\text{реальный}} = \frac{Y_t^{\text{номинальный}}}{P_t}, \quad (1.1)$$

где Y_t – это временной ряд, соответствующий динамике ВВП, а P_t – это уровень цен, который также называют *дефлятором*. Номинальный ВВП обычно выше реального, потому что инфляция обычно является положительным числом.

Таким образом, в макроэкономическом моделировании большой интерес представляет задача прогнозирования показателя реального ВВП государства.

1.2 Методы расчета валового внутреннего продукта и сроки выхода оценок

Существует три метода расчета валового внутреннего продукта, соответствующих различным стадиям воспроизводства:

- по добавленной стоимости (производственный метод);
- по использованию доходов (по расходам);
- по формированию доходов (по доходам).

По всем этим методам мы получаем одно и то же значение ВВП, что следует из *фундаментального тождества счетов национального дохода* [6]

$$\text{совокупный объем производства} = \text{совокупный доход} = \text{совокупный расход}. \quad (1.2)$$

1.2.1 ВВП по добавленной стоимости

В методе расчёта по добавленной стоимости для подсчёта ВВП берутся только конечные товары и услуги, чтобы исключить задвоение в расчётах и завышения ВВП. Промежуточные товары, являющиеся сырьём для производства конечной продукции, во внимание не берутся.

Таким образом, ВВП формируется как сумма валовой добавленной стоимости всех отраслей или институциональных секторов в основных ценах и чистых налогов на продукты за вычетом субсидий [5, 6]

$$Y_t = \sum_{l=1}^L \text{ВДС}_t^{(l)} + \text{НП}_t - \text{СП}_t, \quad (1.3)$$

где

- L – количество всех отраслей (список отраслей приводится в подпараграфе 1.3.1);
- Y_t – это показатель ВВП;
- $\text{ВДС}_t^{(l)}$ – это валовая добавленная стоимость по l -ой отрасли;
- НП_t – это налоги на продукты;
- СП_t – это субсидии на продукты.

Под добавленной стоимостью понимается разность между доходом от продажи выпускаемых товаров и услуг и расходами на закупку материалов и услуг, которые использовались при производстве.

1.2.2 ВВП по использованию доходов

В методе расчёта по использованию доходов суммируются расходы всех экономических субъектов на приобретение конечной продукции. ВВП, рассчитанный по такому методу, формируется как сумма расходов всех институциональных секторов на конечное потребление, валовое накопление, экспорт и импорт [5–7]

$$Y_t = C_t + G_t + N_t + I_t + X_t - M_t, \quad (1.4)$$

где

- Y_t – это показатель ВВП;
- C_t – это расходы на конечное потребление домашних хозяйств;

- G_t – это расходы на конечное потребление государственных организаций;
- N_t – это расходы на конечное потребление частных некоммерческих организаций, обслуживающих домашние хозяйства;
- I_t – это валовое накопление капитала, – инвестиции в основной капитал (оборудование, сооружения, здания и т. д.), в товарные запасы (неиспользованное сырье, незавершенное производство, нереализованная готовая продукция), в жилье (строительство новых жилых домов), а также изменение ценностей (драгоценных металлов и камней, антикварных изделий и т. д.) и природных активов (земли, недр и т. д.);
- X_t – это экспорт, то есть стоимость товаров, произведенных внутри страны и проданных другим странам;
- M_t – это импорт, то есть стоимость товаров иностранного производства, приобретенных отечественными экономическими агентами (домохозяйствами, фирмами, государством).

Нередко при расчете ВВП разницу между экспортом и импортом обозначают

$$E_t = X_t - M_t \quad (1.5)$$

и называют *чистым экспортом*.

Можно объединить две компоненты C_t и N_t в одну в силу экономического смысла, поскольку оба вида расходов связаны с конечным потреблением домашних хозяйств, а также в силу того, что в экономике Беларуси доля вклада компоненты N_t сильно меньше, чем остальных компонент. Таким образом, под обозначением C_t мы будем понимать расходы на конечное потребление домашних хозяйств и частных некоммерческих организаций, обслуживающих домашние хозяйства.

Ввиду сделанных объединений, получаем итоговую формулу метода

$$Y_t = C_t + G_t + I_t + E_t. \quad (1.6)$$

1.2.3 ВВП по доходам

Метод расчёта ВВП по доходам основан на суммировании доходов всех субъектов, которые осуществляют свою деятельность на территории страны, валовой прибыли всех отраслей или институциональных секторов и чистых

налогов на производство и импорт. Подсчет ВВП по такому методу можно выразить в виде формулы [6, 7, 17]

$$Y_t = W_t + N_t + P_t, \quad (1.7)$$

где

- Y_t – это показатель ВВП;
- W_t – это оплата труда или доход работников, то есть заработная плата, премии и так далее;
- N_t – это чистые налоги на производство и импорт;
- P_t – это валовая прибыль и валовые смешанные доходы всех организаций.

1.2.4 Календарь выхода официальных значений ВВП

Ежемесячная оценка ВВП разрабатывается производственным методом в целях получения оперативных данных о текущих изменениях в экономике.

Квартальная оценка ВВП разрабатывается производственным методом и методом использования доходов в целях обеспечения своевременной и более полной информацией о текущей экономической ситуации и установления причин изменений в экономике.

Годовая оценка ВВП является основной. С учетом внутригодовой сезонности важнейших экономических процессов этой оценкой достигается наибольшая согласованность и детализация показателей производства товаров и услуг по различным видам экономической деятельности, использования товаров и услуг на потребление и накопление, формирования доходов.

В таблице 1.1 представлен порядок выхода официальных значений в соответствии с Регламентом оценок ВВП Республики Беларусь [16]. Эти значения публикуются Национальным Статистическим Комитетом (НСК) Республики Беларусь.

Официальные значения ВВП, которые публикует НСК, впоследствии уточняются и величина этих пересмотров может оказаться значительной, что видно из таблицы. В связи с этим оценка динамики ВВП в текущем квартале представляет самостоятельный исследовательский вопрос в теме прогнозирования. Вид прогнозирования ВВП на предыдущий, текущий и следующий квартал в эконометрическом анализе принято называть *наукастингом*.

Таблица 1.1: Регламент выхода официальных значений ВВП

	ВВП производственным методом	ВВП методом конечного использования	ВВП по источникам доходов
Ежемесячные данные (первая оценка)	на 18-20 день месяца, следующего за отчетным	X	X
Квартальные данные (вторая оценка)	на 90 день после квартального периода	на 90 день после квартального периода	на 90 день после квартального периода
Квартальные данные (третья оценка)	декабрь года, следующего за отчетным	декабрь года, следующего за отчетным	декабрь года, следующего за отчетным
Годовые данные (первая оценка)	на 18-20 день месяца, следующего за отчетным годом	на 90 день после отчетного года	на 90 день после отчетного года
Годовые данные (вторая оценка)	декабрь года, следующего за отчетным	декабрь года, следующего за отчетным	декабрь года, следующего за отчетным

1.3 Доли вклада компонент в ВВП Беларуси

Для дальнейшего подбора макроэкономических показателей, по которым мы будем производить прогноз и оценку ВВП, нам необходимо выяснить, какую долю ВВП составляют конкретные компоненты в каждом из трех подходов. Для этого мы рассмотрим соответствующие статистики по Республике Беларусь в период с 2020-го по 2024-ый год.

1.3.1 Доля компонент по добавленной стоимости

Статистику по ВВП в текущих ценах НСК предоставляет по валовой добавленной стоимости 19 отраслей.

Взяв суммарную прибыль за последние 5 лет каждой из отраслей и рассчитав в процентах доли вклада в ВВП, мы можем построить соответствующую диаграмму, представленную на рисунке 1.1. Для уменьшения размеров диаграммы такие отрасли, как «горнодобывающая промышленность», «обра-

батывающая промышленность», «снабжение электроэнергией, газом, паром, горячей водой и кондиционированным воздухом», «водоснабжение; сбор, обработка и удаление отходов, деятельность по ликвидации загрязнений» были объединены в «промышленное производство».



Рисунок 1.1 — Диаграмма долей вклада отраслей в добавленную стоимость

Таким образом, наибольший вклад (более 5%) среди всех отраслей вносят

- промышленное производство;
- оптовая и розничная торговля;
- сельское, лесное и рыбное хозяйство;
- информация и связь;
- операции с недвижимым имуществом;
- строительство;

- транспортная деятельность, складирование, почтовая и курьерская деятельность.

1.3.2 Доля компонент по использованию доходов

В соответствии с формулой (1.6) будем рассматривать данные от НСК по

- расходам на конечное потребление
 - домашних хозяйств и некоммерческих организаций, обслуживающих домашние хозяйства, C_t ;
 - государственных организаций, G_t ;
- валовому накоплению, I_t ;
- чистому экспорту, E_t .

Используя данные по годам за последние 5 лет, вычислим доли по компонентам ВВП

$$d_t^{\frac{C}{Y}} = \frac{C_t}{Y_t}, \quad d_t^{\frac{G}{Y}} = \frac{G_t}{Y_t}, \quad d_t^{\frac{I}{Y}} = \frac{I_t}{Y_t}, \quad d_t^{\frac{E}{Y}} = \frac{E_t}{Y_t}. \quad (1.8)$$

Соответствующие гистограммы за 5 лет представлены на рисунке 1.2.

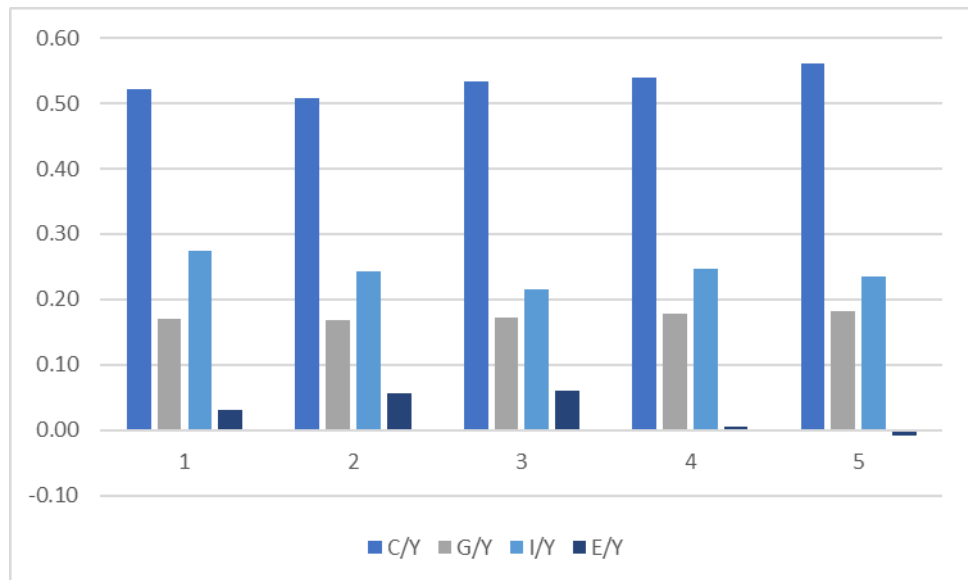


Рисунок 1.2 — Гистограмма долей компонент ВВП по использованию доходов за последние 5 лет

Из гистограммы долей можно заключить, что около 50% от ВВП составляют расходы на конечное потребление домашних хозяйств. Более 20% составляют расходы на валовое накопление. И более 10% составляют расходы на конечное потребление государственных организаций.

1.3.3 Доля компонент по доходам

Данные ВВП по источникам доходов в текущих ценах НСК предоставляет по следующим компонентам

- оплата труда работников, W_t ;
- чистые налоги на производство и импорт N_t ;
- валовая прибыль и валовые смешанные доходы организаций P_t .

Вычислим доли по компонентам ВВП

$$d_t^{\frac{W}{Y}} = \frac{W_t}{Y_t}, \quad d_t^{\frac{N}{Y}} = \frac{N_t}{Y_t}, \quad d_t^{\frac{P}{Y}} = \frac{P_t}{Y_t}. \quad (1.9)$$

На рисунке 1.3 представлены соответствующие гистограммы за 5 лет.

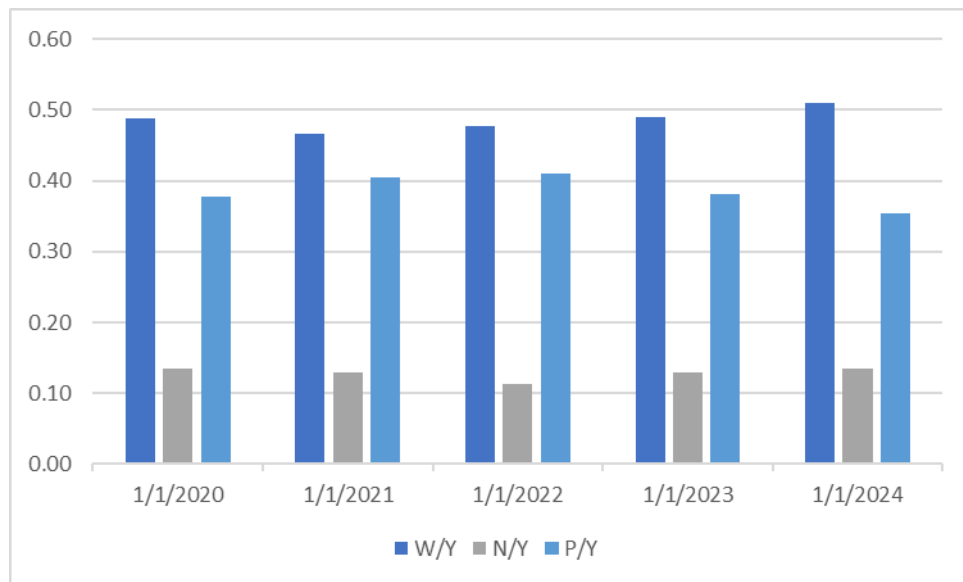


Рисунок 1.3 — Гистограмма долей компонент ВВП по источникам доходов за последние 5 лет

Следовательно, около 50% от ВВП составляет оплата труда работников, около 40% – валовые доходы организаций и более 10% – чистые налоги на производство и импорт.

Таким образом, при построении макроэкономической модели реального ВВП с экономической точки зрения нас больше интересуют именно те компоненты, которые вносят наибольший вклад в ВВП.

1.4 Опыт краткосрочного прогнозирования и наукастинга реального ВВП на основе моделей по данным смешанной частоты

Оценки ВВП формируются НСК Республики Беларусь на квартальной и годовой частоте. В то же время отраслевые показатели, по которым формируется оценка ВВП, публикуются на месячной частоте. В связи с этим становится актуальным вопрос об использовании более оперативной информации при прогнозировании показателя ВВП. Эта задача называется *задачей макроэкономического моделирования реального ВВП по данным смешанной частоты*. Решению этой задачи посвящена данная работа.

Для решения задачи макроэкономического моделирования показателя ВВП на квартальной частоте по оперативным показателям высокой частоты наиболее популярными одномерными подходами являются:

- построение авторегрессионных моделей с агрегированием данных высокой частоты к низкой частоте;
- построение авторегрессионных моделей с интерполированием данных низкой частоты к высокой частоте;
- построение авторегрессионных моделей MIDAS по данным смешанной частоты.

1.4.1 Авторегрессионные модели с агрегацией и интерполяцией

Авторегрессионные модели проинтегрированного скользящего среднего активно используются в эконометрическом анализе для прогнозирования, поскольку позволяют учитывать автокорреляцию и тренды в данных. Исследования по США свидетельствуют, что модель AR(2) может превосходить сложные модели в периоды стабильности, но теряет эффективность при кризисах.

Приведем самое общее представление для авторегрессионных моделей проинтегрированного скользящего среднего – SARIMAX (*Seasonal AutoRegressive Integrated Moving Average with eXogenous regressors*). Формально модель SARIMAX(p, d, q, P, D, Q) можно представить в виде

$$\begin{aligned} \Delta^d \Delta_s^D y_t = & \beta_0 + \sum_{j=1}^p \beta_j \Delta^d \Delta_s^D y_{t-j} + \varepsilon_t + \sum_{i=1}^q \alpha_i \varepsilon_{t-i} + \\ & + \sum_{J=1}^P \delta_J \Delta^d \Delta_s^D y_{t-Js} + \sum_{I=1}^Q \gamma_I \varepsilon_{t-Is} + \sum_{k=1}^n \zeta_k x_{k,t}, \end{aligned} \quad (1.10)$$

где

- $y_t, t = 1, \dots, T$ – эндогенный временной ряд;
- $x_{1,t}, \dots, x_{n,t} - n$ экзогенных временных рядов;
- $\varepsilon_t \sim WN(0, \sigma^2)$ – белый шум;
- p, d, q – порядки авторегрессии, интегрирования и скользящего среднего соответственно;
- P, D, Q – сезонные порядки авторегрессии, интегрирования и скользящего среднего соответственно;
- s – период сезонности;
- Δ^d – оператор конечной разности порядка d ;
- Δ_s^D – оператор конечной разности порядка D с периодом s ;
- $\alpha_i, \beta_j, \gamma_I, \delta_J, \zeta_k \in \mathbb{R}$.

Модель SARIMAX позволяет моделировать поведение временного ряда y_t , с единичным корнем порядка d и сезонной компонентой, используя n экзогенных временных рядов той же частоты $x_{k,t}$. То есть эта модель определяется для временных рядов одинаковой частоты. Для того, чтобы она работала для временных рядов смешанной частоты, необходимо свести все временные ряды к одной частоте.

Агрегирование временных рядов высокой частоты к низкой частоте. В зависимости от экономического смысла временного ряда, можно проводить агрегирование с высокой частоты к низкой частоте по среднему или суммированию. В частности, для показателей запаса агрегированное значение на данный момент времени определяется как среднее арифметическое значений показателя за период; для показателей потока агрегированное значение на данный момент времени определяется как сумма значений показателя за период.

Преимущества такого подхода заключаются в простоте его реализации. Однако вместе с этим происходит потеря некоторой информации о внутривременной динамике высокочастотных рядов, из-за чего оценки рискуют оказаться смещенными.

Интерполирование временных рядов низкой частоты к высокой частоте. В некоторых случаях можно применить методы полиномиальной интерполяции (например, кубические сплайны) или методы на основе структурных моделей (например, метод Chow-Lin [40]).

Основным недостатком такого подхода служит появление искусственной волатильности и ошибок спецификации. То есть появляется некоторая информация, которая изначально не была заложена в поведение исходного временного ряда низкой частоты.

Таким образом, SARIMAX модели с агрегированием или интерполированием данных приводят либо к потере некоторой информации о высокочастотных рядах, либо к появлению искусственной, возможно ложной, информации о низкочастотных рядах.

1.4.2 Регрессионные MIDAS модели

Альтернативным подходом для работы с данными смешанной частоты являются регрессионные модели MIDAS (*MIxed DAta Sampling*) [23, 27, 28]. Пусть временной ряд y_t , $t = 1, \dots, T$ является эндогенной переменной и задан на фиксированной частоте. Кроме того, пусть временные ряды $x_t^{(m_1)}, \dots, x_t^{(m_n)}$ являются экзогенными переменными, которые заданы в m_j раз чаще. То есть у каждого временного ряда $x^{(m_j)}$, вообще говоря, своя частота.

Функция лаговых коэффициентов

$$b(i, \Theta), \quad i = 0, \dots, p. \quad (1.11)$$

играет решающую роль в моделях по данным смешанной частоты, так как она выступает в роли «агрегатора» данных высокой частоты к данным низкой частоты. Различные варианты задания этой функций (например, Almon или Beta полиномы [22]) будут по-разному справляться с решением задачи прогнозирования. Используя функцию лаговых коэффициентов, классическую многофакторную модель MIDAS по данным смешанной частоты можно формально записать в виде

$$y_t = \beta_0 + \sum_{j=0}^n \beta_j \sum_{i=0}^p b(i, \Theta) L^{i/m_j} x_t^{(m_j)} + \varepsilon_t, \quad (1.12)$$

где β_0, β_1 – это некоторые вещественные коэффициенты.

Одним из распространенных вариантов моделей MIDAS регрессии является модель U-MIDAS (*Unrestricted MIXed Data Sampling*) [31]. В отличие от базовой модели MIDAS, в U-MIDAS не накладываются ограничения на полиномиальный лаговый оператор, то есть такой подход функции лаговых коэффициентов (1.11) заменяются на постоянные значения.

В общем виде U-MIDAS модель с одной экзогенной переменной может быть записана в следующем виде

$$y_t = \beta_0 + \sum_{j=1}^m \sum_{i=0}^p \beta_{i,j} x_{(t-i),j} + \varepsilon_t. \quad (1.13)$$

То есть временной ряд x_t , заданный в m раз чаще, чем y_t , раскладывается в m рядов той же частоты, что и y_t .

Ключевым преимуществом базовой MIDAS модели является то, что она допускает длинные лаги при ограниченном количестве параметров, что может быть особенно полезно в финансовых приложениях с большим несоответствием между частотами рядов y_t и x_t . Например, когда y_t является ежемесячным показателем, а x_t — ежедневным.

Однако для макроэкономических приложений с небольшими различиями в частотах выборки, например, для ежемесячных и квартальных данных, базовая модель может иметь определенные недостатки. Например, базовая модель сильно нелинейна по параметрам, так что она не может быть оценена с помощью метода наименьших квадратов, в отличие от U-MIDAS. В целом предполагается, что модель U-MIDAS должна работать лучше, чем базовая модель MIDAS, пока частота агрегации мала, а также пока U-MIDAS не слишком сильно параметризована.

Таким образом, используя MIDAS и U-MIDAS подходы можно строить эконометрические модели временных рядов по данным смешанной частоты. Данные подходы активно используются в одномерном случае, когда моделируется поведение одного экономического показателя на низкой частоте, зависящего от одного или нескольких показателей на высокой частоте.

1.4.3 Обзор полученных результатов в исследованиях о наукастинге ВВП

Стратегии макроэкономического моделирования и наукастинга ВВП основаны на использовании дополнительной информации из опережающих (ежемесячных, еженедельных, ежедневных) показателей, отражающих текущее состояние экономики. В частности, таковыми являются некоторые ком-

поненты, входящие в основные методы расчета ВВП.

В самых ранних исследованиях для наукастинга применялась модель множественной регрессии, в которой рост ВВП за квартал предсказывался на основе сбалансированной панели ежемесечных переменных, агрегированных в квартальное выражение [37, 38, 41, 42]. Однако для проведения наукастинга требуются все значения ежемесечных наблюдений за прогнозируемый квартал. На практике эти значения прогнозировались с помощью отдельных моделей. Помимо этого, как ранее было указано, данный способ приводит к потере информации из высокочастотных переменных.

В более поздних исследованиях применяется модель MIDAS. На примере задачи наукастинга ВВП в США и еврозоне эксперименты Foroni и др. [31] показали, что модели U-MIDAS не уступают в точности традиционным моделям MIDAS с экспоненциальными лагами Almon. В своей работе Foroni and Macellino [29] по наукастингу ВВП еврозоны выяснили, что рассматриваемые модели MIDAS регрессии превосходят по точности модели AR с агрегацией.

В приложениях к экономике России Макеева и Станкевич [4] показывают, что наукастинг годовых темпов роста ВВП России по U-MIDAS и MIDAS оказывается стабильно точнее, чем по динамическим факторным моделям и AR(1). А MIDAS с марковскими переключениями режимов, оказываются наиболее точными в сравнении с традиционными моделями MIDAS [14].

Кроме одномерных регрессионных моделей для решения задачи наукастинга ВВП также применяются модели многомерных временных рядов. Среди многомерных моделей одним из наиболее распространенных подходов к наукастингу ВВП является построение векторной авторегрессии по данным смешанной частоты (*Mixed Frequency Vector AutoRegression*, MF-VAR). Модели VAR хорошо себя показали в работах по наукастингу ВВП России. В частности, в работе Станкевича [15] проводится наукастинг с помощью MIDAS и MF-VAR для разных тестовых выборок. В итоге самую высокую точность демонстрируют именно модели векторной авторегрессии. А в исследованиях Макеевой и Станкевича [4] прогноз темпов роста ВВП по MF-VAR оказывается точнее, чем в DFM, MIDAS и AR(1).

Таким образом, в данной работе особый интерес представляют именно многомерные модели MF-VAR для макроэкономического моделирования и наукастинга реального ВВП Беларуси.

ГЛАВА 2

МНОГОМЕРНЫЕ МОДЕЛИ И МЕТОДЫ АНАЛИЗА ПО ДАННЫМ СМЕШАННОЙ ЧАСТОТЫ

В настоящей главе приводится формальное описание моделей, используемых для решения задачи прогнозирования. Приводятся классификации моделей векторных авторегрессий и их свойства. Объясняется метод оценки точности построенных моделей.

2.1 Модель векторной авторегрессии VAR и ее применение в задачах макроэкономического анализа

Векторная авторегрессия обычно используется в эконометрике для прогнозирования систем со взаимосвязанными временными рядами и для анализа динамического воздействия случайных возмущений на систему переменных. Данный подход позволяет рассматривать каждую эндогенную переменную в системе как функцию от всех эндогенных переменных в системе с их лагами до порядка p .

2.1.1 Определение модели векторной авторегрессии

Пусть задан вектор временных рядов

$$\mathbf{y}_t = \begin{pmatrix} y_{t,1} \\ \vdots \\ y_{t,n} \end{pmatrix}, \quad t = 1, \dots, T,$$

где $y_{i,t}$, $i = 1, \dots, n$ – это i -ый временной ряд, соответствующий конкретному макроэкономическому показателю, t – время. Введем векторные параметры

$$\boldsymbol{\beta}_0 = \begin{pmatrix} \beta_{0,1} \\ \vdots \\ \beta_{0,n} \end{pmatrix}, \quad \beta_{0,i} \in \mathbb{R}, \quad i = 1, \dots, n,$$

$$\mathbf{B}_j = \begin{pmatrix} \beta_{j,11} & \dots & \beta_{j,1n} \\ \vdots & \ddots & \vdots \\ \beta_{j,n1} & \dots & \beta_{j,nn} \end{pmatrix}, \quad \beta_{j,ik} \in \mathbb{R}, \quad i, k = 1, \dots, n, \quad \forall j$$

и вектор шумов

$$\boldsymbol{\varepsilon}_t = \begin{pmatrix} \varepsilon_{t,1} \\ \vdots \\ \varepsilon_{t,n} \end{pmatrix}, \quad \varepsilon_{t,i} \sim WN(0, \sigma^2), \quad i = 1, \dots, n,$$

где $WN(0, \sigma^2)$ – это белый шум со средним 0 и дисперсией σ^2 . Тогда модель *векторной авторегрессии порядка p* (*Vector AutoRegression*, VAR) можно записать в виде

$$\mathbf{y}_t = \boldsymbol{\beta}_0 + \sum_{j=1}^p \mathbf{B}_j \mathbf{y}_{t-j} + \boldsymbol{\varepsilon}_t. \quad (2.1)$$

Причем для вектора шоков отсутствует серийная корреляция

$$\text{corr}(\varepsilon_{t,i}, \varepsilon_{s,i}) = 0, \quad t \neq s,$$

и допускается кросс-корреляция

$$\text{corr}(\varepsilon_{t,i}, \varepsilon_{t,j}), \quad i \neq j.$$

Если в уравнении (2.1) обозначить

$$\mathbf{B}(z) = \mathbf{E} - z\mathbf{B}_1 - \dots - z^p\mathbf{B}_p,$$

где \mathbf{E} – это единичная матрица размерности $n \times n$, то критерий стационарности авторегрессионной модели для многомерного случая определяется следующим образом: все корни уравнения

$$\det \mathbf{B}(z) = 0 \quad (2.2)$$

должны быть по модулю больше единицы.

Для определения количества лагов p применяются многомерные аналоги информационных критериев Акаике, Шварца, Хеннана-Куина, а для оценки коэффициентов авторегрессии \mathbf{B}_j , $j = 1, \dots, p$ применяется метод наименьших квадратов.

2.1.2 Причинность по Грейнджеру

Эмпирический анализ часто затрудняет выявление причинно-следственных связей. Даже при наличии сильной корреляции между переменными, без дополнительной информации невозможно точно установить причины их взаимного влияния. Иногда экономическая теория позволяет установить взаимосвязь между переменными. Но в случаях, когда экономической теории недостаточно или же она дает неоднозначный ответ, VAR модель дает возможность определить причинность переменных, то есть обнаружить как изменение одних переменных предшествует изменению других, основываясь только на статистических результатах

Если рассматривать два фактора x и y , то под причинностью понимается следующее: если переменная x влияет на переменную y , то это означает, что y меняется либо мгновенно, либо через короткое время после изменения x . Или наоборот, если x не является причиной изменения y , то изменения x не влияют на будущие значения y .

Пусть $\{y_t\}$, $\{x_t\}$ – временные ряды, $t \in T$. Построим для них VAR(p) модель

$$\begin{aligned}x_t &= \mu_1 + \sum_{j=1}^p \alpha_j x_{t-j} + \sum_{j=1}^p \beta_j y_{t-j} + u_t, \\y_t &= \mu_2 + \sum_{j=1}^p \gamma_j x_{t-j} + \sum_{j=1}^p \delta_j y_{t-j} + v_t,\end{aligned}$$

где $u_t, v_t \sim WN(0, \sigma^2)$, а $\alpha_j, \beta_j, \gamma_j, \delta_j \in \mathbb{R}$ – это некоторые коэффициенты, которые определяются по методу наименьших квадратов. Для этой модели тестируются две гипотезы

$$\begin{aligned}H'_0 &: \beta_1 = \dots = \beta_p = 0, \\H''_0 &: \gamma_1 = \dots = \gamma_p = 0.\end{aligned}$$

Тогда x_t является причиной по Грейнджеру для y_t , если H'_0 не отвергается, а H''_0 отвергается. И наоборот, y_t является причиной по Грейнджеру для x_t , если H''_0 не отвергается, а H'_0 отвергается.

Важно понимать, что причинность по Гренджеру не означает наличие причинно-следственной связи, но отсутствие причинности по Гренджеру подтверждает отсутствие причинно-следственной связи.

2.1.3 Функция импульсных откликов

В общем случае коэффициенты VAR модели не имеют экономической интерпретации. Вследствие этого принято рассматривать *функции импульсных откликов* (impulse response function, IRF) как описание влияния экзогенного шока на показатели. Функция импульсного отклика задает траекторию по h , которую описывает y_{t+h} , инициированная единичным шоком ε_t .

Сперва рассмотрим модель AR(1)

$$y_t = \beta y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2), \quad \beta \in \mathbb{R}. \quad (2.3)$$

Ее можно представить в виде модели скользящего среднего порядка $+\infty$

$$y_t = \beta y_{t-1} + \varepsilon_t = \beta \cdot (\beta y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \dots = \sum_{j=0}^{+\infty} \beta^j \varepsilon_{t-j}. \quad (2.4)$$

Тогда из этого представления видно, что если задать единичный шок

$$\varepsilon_0 = 0, \quad \varepsilon_1 = 1, \quad \varepsilon_2 = 0, \dots, \varepsilon_j = 0, \dots$$

то в соответствии с выражением (2.4)

$$\begin{aligned} y_0 &= \varepsilon_0 = 0, \\ y_1 &= \varepsilon_1 + \beta \varepsilon_0 = 1, \\ y_2 &= \varepsilon_2 + \beta \varepsilon_1 + \beta^2 \varepsilon_0 = \beta, \\ &\vdots \\ y_j &= \beta^j, \end{aligned}$$

и так далее. Таким образом, для одномерной модели AR(1) функция импульсного отклика – это то же самое, что и представление в виде модели скользящего среднего порядка $+\infty$.

Формально функция импульсного отклика для VAR модели – это матричная функция от $h = 0, 1, 2, \dots$, которая задается как

$$\frac{\partial \mathbf{y}_{t+h}}{\partial \varepsilon_t} = \left(\frac{\partial y_{i,t+h}}{\partial \varepsilon_{j,t}} \right)_{p \times p} \quad (2.5)$$

Таким образом, элемент (i, j) этой матричной функции отражает реакцию переменной y_i на единичное изменение шока в переменной y_j в момент времени $t + h$. В итоге функция импульсного отклика для конкретной переменной y_i

на шок в переменной y_j задается последовательностью

$$\frac{\partial y_{i,t}}{\partial \varepsilon_{j,t}}, \frac{\partial y_{i,t+1}}{\partial \varepsilon_{j,t}}, \dots, \frac{\partial y_{i,t+h}}{\partial \varepsilon_{j,t}}, \dots$$

Рассмотрим модель VAR(1)

$$\mathbf{y}_t = \mathbf{B}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \quad (2.6)$$

в соответствии с обозначениями (2.1). Эту модель также можно представить в виде векторной модели скользящего среднего

$$\mathbf{y}_t = \mathbf{B}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t = \mathbf{B}(\mathbf{B}\mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_{t-1}) + \boldsymbol{\varepsilon}_t = \dots = \sum_{j=0}^{\infty} \mathbf{B}^j \boldsymbol{\varepsilon}_{t-j}.$$

Тогда

$$\mathbf{y}_{t+h} = \sum_{j=0}^{\infty} \mathbf{B}^j \boldsymbol{\varepsilon}_{t+h-j}, \quad (2.7)$$

откуда

$$\frac{\partial y_{t+h}}{\partial \varepsilon_{j,t}} = \frac{\partial}{\partial \varepsilon_{j,t}} \left(\sum_{j=0}^{\infty} \mathbf{B}^j \boldsymbol{\varepsilon}_{t+h-j} \right) = \mathbf{B}^h e_j, \quad (2.8)$$

где e_j – это j -ый столбец единичной матрицы размерности $p \times p$. Таким образом, элемент $\beta_{ij,h}$ в матрице \mathbf{B}^h описывает отклик переменной $y_{i,t}$ через h шагов в будущем на единичный шок $\varepsilon_{j,t}$.

Возможны случаи, когда шоки $\boldsymbol{\varepsilon}_t$ коррелированы. Тогда мы не можем рассуждать о рассматриваемых шоках как об изолированных, поскольку в силу корреляции будет наблюдаться мгновенное воздействие на все другие элементы вектора. Поэтому, с учетом этих единовременных эффектов, рассматриваются ортогонализированные функции импульсного отклика, например, через разложение Холецкого.

Пусть $\boldsymbol{\Sigma}$ – это ковариационная матрица ошибок. Тогда, используя разложение Холецкого, ее можно представить в виде

$$\boldsymbol{\Sigma} = \mathbf{P} \cdot \mathbf{P}^T. \quad (2.9)$$

Тогда из уравнения (2.7) получим

$$\mathbf{y}_{t+h} = \sum_{j=0}^{\infty} \mathbf{B}^j \mathbf{P} \mathbf{P}^{-1} \boldsymbol{\varepsilon}_{t+h-j} = \sum_{j=0}^{\infty} \mathbf{B}^j \mathbf{P} \mathbf{v}_{t+h-j}, \quad (2.10)$$

где

$$\mathbf{v}_t = \mathbf{P}^{-1} \boldsymbol{\varepsilon}_t. \quad (2.11)$$

В силу этой замены получаем, что

$$\boldsymbol{\varepsilon}_t = \mathbf{P} \mathbf{v}_t, \quad (2.12)$$

то есть исходные ошибки будут заданы как

$$\begin{aligned} \varepsilon_{t,1} &= p_{11}v_{t,1}, \\ \varepsilon_{t,2} &= p_{21}v_{t,1} + p_{22}v_{t,2}, \\ &\vdots \\ \varepsilon_{t,n} &= p_{n1}v_{t,1} + \dots + p_{nn}v_{t,n}. \end{aligned}$$

Причем новые ошибки v_t не коррелируют друг с другом

$$\mathbb{E}\{\mathbf{v}_t \mathbf{v}_t^T\} = \mathbf{P}^{-1} \mathbb{E}\{\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t^T\} (\mathbf{P}^{-1})^T = \mathbf{P}^{-1} \boldsymbol{\Sigma} (\mathbf{P}^{-1})^T = \mathbf{P}^{-1} \mathbf{P} \mathbf{P}^T (\mathbf{P}^{-1})^T = \mathbf{E}.$$

Таким образом, мы получим

$$\frac{\partial y_{t+h}}{\partial \varepsilon_{j,t}} = \mathbf{B}^h \mathbf{P} e_j, \quad (2.13)$$

то есть матрицу импульсных откликов можно строить как в случае некоррелированных ошибок, так и в случае коррелированных.

Также заметим, что если временные ряды в VAR модели заданы в процентах, то в случае единичного шока с помощью функции импульсных откликов можно наблюдать, на сколько процентов будут изменяться остальные переменные, а тогда сами отклики будут совпадать с эластичностям между двумя рядами в момент времени h .

Таким образом, для модели VAR мы определили два важных свойства: причинность по Грейнджеру и функции импульсных откликов. По наличию причинности по Грейнджеру и значимых откликов можно утверждать об опережающем характере одних переменных по отношению к другим.

2.2 Модель векторной авторегрессии по данным смешанной частоты MF-VAR

Как мы определили ранее, стандартная модель VAR требует, чтобы все переменные имели одинаковую частоту. Этого можно добиться путем агрегирования данных высокой частоты, но в таком случае возникает потеря информации. В одномерных моделях потеря информации устраняется путем применения модели MIDAS регрессии, которая на основе высокочастотных переменных создает новые низкочастотные переменные, которые сохраняют точность. Однако такой подход может привести к резкому увеличению числа нелинейных параметров, если использовать полиномиальные лаговые веса, особенно в случае многих экзогенных переменных. Чтобы уменьшить количество параметров до более управляемого уровня, в тех случаях, когда количество добавляемых переменных невелико, используется подход U-MIDAS.

Под моделью MF-VAR (*Mixed Frequency Vector AutoRegression*) понимается такая VAR модель, для которой данные высокой частоты агрегируются с помощью модели U-MIDAS (1.13). То есть временной ряд высокой частоты, который замерен в m раз чаще, чем ряд низкой частоты, раскладывается на m временных рядов низкой частоты.

Пусть задан временной ряд $\{y_t^Q\}$ на квартальной частоте и $\{x_t^M\}$ на месячной частоте, $t \in T$. В соответствии с U-MIDAS подходом, раскладывая временной ряд $\{x_t^M\}$ месячной частоты на три временных ряда квартальной частоты $\{x_t^{M_1}\}$, $\{x_t^{M_2}\}$, $\{x_t^{M_3}\}$, мы можем построить модель MF-VAR(p) как VAR(p) вида

$$\begin{aligned} y_t^Q &= \mu_1 + \sum_{j=1}^p \beta_j^{1,1} x_{t-j}^{M_1} + \sum_{j=1}^p \beta_j^{1,2} x_{t-j}^{M_2} + \sum_{j=1}^p \beta_j^{1,3} x_{t-j}^{M_3} + \sum_{j=1}^p \beta_j^{1,4} y_{t-j}^Q + \varepsilon_{t,1}^Q, \\ x_t^{M_1} &= \mu_2 + \sum_{j=1}^p \beta_j^{2,1} x_{t-j}^{M_1} + \sum_{j=1}^p \beta_j^{2,2} x_{t-j}^{M_2} + \sum_{j=1}^p \beta_j^{2,3} x_{t-j}^{M_3} + \sum_{j=1}^p \beta_j^{2,4} y_{t-j}^Q + \varepsilon_{t,2}^Q, \\ x_t^{M_2} &= \mu_3 + \sum_{j=1}^p \beta_j^{3,1} x_{t-j}^{M_1} + \sum_{j=1}^p \beta_j^{3,2} x_{t-j}^{M_2} + \sum_{j=1}^p \beta_j^{3,3} x_{t-j}^{M_3} + \sum_{j=1}^p \beta_j^{3,4} y_{t-j}^Q + \varepsilon_{t,3}^Q, \\ x_t^{M_3} &= \mu_4 + \sum_{j=1}^p \beta_j^{4,1} x_{t-j}^{M_1} + \sum_{j=1}^p \beta_j^{4,2} x_{t-j}^{M_2} + \sum_{j=1}^p \beta_j^{4,3} x_{t-j}^{M_3} + \sum_{j=1}^p \beta_j^{4,4} y_{t-j}^Q + \varepsilon_{t,4}^Q, \end{aligned} \tag{2.14}$$

где $\mu_i \in \mathbb{R}$, $\varepsilon_{t,i}^Q \sim WN(0, \sigma^2)$, $i = \overline{1,4}$.

Обозначим векторы

$$\mathbf{Y}_t^Q = \begin{bmatrix} y_t^Q \\ x_{t,1}^Q \\ x_{t,2}^Q \\ x_{t,2}^Q \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{bmatrix}, \quad \mathbf{B}_j = \begin{bmatrix} \beta_j^{1,1} & \beta_j^{1,2} & \beta_j^{1,3} & \beta_j^{1,4} \\ \beta_j^{2,1} & \beta_j^{2,2} & \beta_j^{2,3} & \beta_j^{2,4} \\ \beta_j^{3,1} & \beta_j^{3,2} & \beta_j^{3,3} & \beta_j^{3,4} \\ \beta_j^{4,1} & \beta_j^{4,2} & \beta_j^{4,3} & \beta_j^{4,4} \end{bmatrix}, \quad \boldsymbol{\varepsilon}_t^Q = \begin{bmatrix} \varepsilon_{t,1}^Q \\ \varepsilon_{t,2}^Q \\ \varepsilon_{t,3}^Q \\ \varepsilon_{t,4}^Q \end{bmatrix}.$$

и тогда можно записать MF-VAR(p) модель в виде

$$\mathbf{Y}_t^Q = \boldsymbol{\mu} + \sum_{j=1}^p \mathbf{B}_j \times \mathbf{Y}_{t-j}^Q + \boldsymbol{\varepsilon}_t^Q. \quad (2.15)$$

Очевидно, что в модель можно добавлять больше низкочастотных и высокочастотных переменных, но надо учитывать, что высокочастотные переменные будут также раскладываться на три компоненты.

Остановимся подробно на количестве оцениваемых коэффициентов в модели. Если модель содержит n_Q квартальных переменных и n_M месячных переменных, то общее количество уравнений $n_Q + 3n_M$. В каждом уравнении требуется оценить $1 + (n_Q + 3n_M) \times p$ коэффициентов. Тогда для оценки всей системы требуется оценить

$$[1 + (n_Q + 3n_M) \times p] \times (n_Q + 3n_M)$$

коэффициентов. Следовательно, при добавлении в модель новых высокочастотных факторов число параметров будет быстро возрастать. Поэтому значение p должно быть относительно небольшим и высокочастотные переменные требуют тщательного отбора.

Таким образом, мы формально определили модель, которая будет применяться для решения задачи прогнозирования. Проблема выбора количества лагов p и отбора высокочастотных переменных становится актуальной в условиях применения данной модели. Также она сохраняет свойства стандартной VAR модели, такие как причинность по Грейнджеру и функции импульсных откликов.

2.3 Оценка точности модельных прогнозов

Имея определенную модель для исследований необходимо сформулировать характеристики, по которым будет проводиться качественная оценка прогнозов модели.

2.3.1 Критерии оценки точности прогнозов

Для оценки качества прогнозов эконометрических моделей наиболее популярными являются два следующих критерия:

- MAE (*Mean Absolute Error*)

$$MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|; \quad (2.16)$$

- RMSE (*Root Mean Squared Error*)

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2}, \quad (2.17)$$

где

- y_t — фактическое значение эндогенной переменной в период t ;
- \hat{y}_t — спрогнозированное с помощью модели значение эндогенной переменной в период t ;
- T — количество периодов, на которых тестируется модель.

Остается определить период T , на котором будет тестироваться модель.

2.3.2 Принцип наукастинга

При построении наукастов не учитывается информация о последнем доступном квартале: перед оцениванием модели из выборки удаляются значения зависимых переменных и соответствующие данному кварталу месячные значения объясняющих переменных. Далее в выборку возвращаются удаленные значения регрессоров и для них рассчитывается прогнозное значение зависимой переменной (наукаст).

Для проверки устойчивости к добавлению новых данных все модели тестируются трижды: с использованием данных по объясняющим переменным за все три месяца квартала, для которого рассчитывается наукаст; без данных за последний месяц и без данных за два последних месяца. В случае удаления части данных в объясняющих переменных, «пустым» месяцам в объясняющих переменных проставляется последнее доступное значение показателя (за второй или за первый месяц квартала в зависимости от метода тестирования). Такая проверка позволяет определить, насколько методы устойчивы к объему используемых данных, и смоделировать встречающиеся в реальной жизни условия, когда наукаст показателя за текущий квартал рассчитывается еще до окончания квартала.

2.3.3 Алгоритм оценки точности по методу расширяющегося окна

Оценка модели по одному вневыборочному значению может оказаться недостаточно точной. Более точные результаты можно получить, построив оценки по нескольким вневыборочным точкам. Но в силу того, что MF-VAR модель пригодна для краткосрочного прогнозирования, строить прогноз на несколько точек вперед не эффективно.

Для решения описанной проблемы используется подход *расширяющегося окна*.

1. Вся выборка данных первоначально делится на две части – обучающую и тестируемую. В результате в обучающую выборку входят наблюдения y_1, \dots, y_t , а в тестируемую – наблюдения y_{t+1}, \dots, y_T .
2. Основываясь на фактических временных рядах обучающей выборки осуществляется оценка параметров модели.
3. С помощью оцененной модели производится прогноз \hat{y}_{t+1} на один квартал вперед.
4. Увеличивается длина обучающей выборки на один квартал, то есть теперь в нее входят наблюдения y_1, \dots, y_t, y_{t+1} , и заново оцениваются параметры модели.
5. Используя переоцененную модель с учетом увеличенной выборки, производится прогноз \hat{y}_{t+2} на один квартал вперед.
6. Далее аналогично шагам 4 и 5 последовательно увеличивается выборка, переоценивается модель и строится прогноз на один квартал вперед до

тех пор, пока выборка не будет включать все наблюдения y_1, \dots, y_T .

7. В результате получены точечные прогнозы $\hat{y}_{t+1}, \dots, \hat{y}_T$ темпов прироста реального ВВП, которые сравниваются с фактическими данными и рассчитываются метрики качества модели

$$MAE = \frac{1}{T - t - 1} \sum_{i=t+1}^T |y_i - \hat{y}_i|,$$
$$RMSE = \sqrt{\frac{1}{T - t - 1} \sum_{i=t+1}^T (y_i - \hat{y}_i)^2}.$$

Приведенный алгоритм расширения выборки с переоценкой коэффициентов и называется *методом расширяющегося окна*.

Окно можно расширять также на два значения y_{t+1}, y_{t+2} или более. Но тогда могут возникать коллизии с тем, что и на первом, и на втором шаге алгоритма строятся, вообще говоря, разные оценки \hat{y}_{t+2} . В таком случае нужно подробно останавливаться на выборе конкретной оценки для y_{t+2} : как среднего или как значения с наименьшим отклонением.

Таким образом, алгоритм расширяющегося окна позволяет оценивать точность краткосрочных прогнозов на некотором периоде $t + 1, \dots, T$.

ГЛАВА 3

ПРОГНОЗИРОВАНИЕ ТЕМПОВ ПРИРОСТА РЕАЛЬНОГО ВВП НА ОСНОВЕ МОДЕЛИ MF-VAR

В настоящей главе приводится описание макроэкономических данных и задачи исследования. Рассматривается общая схема построения модели с проведением предварительного анализа временных рядов. Выполняется построение моделей на основе MF-VAR и VAR с последующей оценкой точности прогнозов моделей и сравнительным анализом полученных результатов.

3.1 Постановка задачи и описание исследуемых макроэкономических показателей

Национальным Банком Республики Беларусь сформулирована следующая прикладная задача по построению макроэкономической модели белорусской экономики. Необходимо построить математическую модель роста реального квартального ВВП Республики Беларусь по опережающим месячным показателям, провести прогноз с помощью этой модели, оценить результаты этого прогноза и свойства построенной модели.

Для решения задачи прогнозирования реального квартального ВВП по данным месячной частоты Национальным Банком были предложены следующие показатели на месячной частоте:

- денежные доходы населения, базисный индекс объема (янв. 2018 = 1);
- объем инвестиций в основной капитал в среднегодовых ценах 2018 г.;
- объем продукции сельского хозяйства, в среднегодовых ценах 2018 г.;
- объем промышленного производства в среднегодовых ценах 2018 г.;
- объем розничного товарооборота в среднегодовых ценах 1995 г.;
- базисный индекс физического объема строительно-монтажных работ (янв. 2018 = 1);
- сводный индекс экономических настроений.

Подбор показателей обусловлен тем, что оценки по этим показателям обновляются с наименьшей задержкой, то есть они являются *опережающими* для ВВП.

Важно установить характер экономической взаимосвязи в будущей математической модели. Для этого необходимо определить экономический смысл каждого из приведенных показателей и его взаимосвязь с показателем ВВП.

3.1.1 Объем денежных доходов населения

Объем денежных доходов населения – это макроэкономический показатель, который отражает совокупный объем всех денежных средств, полученных населением страны за определенный период времени (обычно месяц, квартал или год). Этот показатель включает доходы от трудовой деятельности, предпринимательства, социальные выплаты (пенсии, пособия, стипендии), доходы от собственности и другие поступления.

Данный показатель является важным индикатором уровня благосостояния населения и экономической активности в стране, так как он напрямую связан с покупательной способностью граждан и уровнем их жизни.

Объем денежных доходов населения является аппроксимацией компоненты W_t в формуле расчета ВВП по доходам (1.7). А из рисунка 1.3 было заключено, что доля этой компоненты в ВВП составляет почти 50%.

3.1.2 Объем инвестиций в основной капитал

Объем инвестиций в основной капитал – это макроэкономический показатель, который отражает совокупный объем денежных средств, направленных на создание, модернизацию, приобретение и ремонт основных средств производства в экономике страны за определенный период времени. По технологической структуре инвестиции в основной капитал подразделяются на следующие виды работ и затрат: строительно-монтажные работы (включая работы по монтажу оборудования); затраты на приобретение машин, оборудования, транспортных средств, инструмента и инвентаря; прочие работы и затраты.

Этот показатель является одним из главных индикаторов инвестиционной активности и экономического развития, так как напрямую связан с формированием производственного потенциала страны.

Объем инвестиций в основной капитал является аппроксимацией компоненты I_t в формуле расчета ВВП по расходам (1.6). Также из рисунка 1.2 было заключено, что доля валового накопления I_t в ВВП составляет более 20%.

3.1.3 Объем продукции сельского хозяйства

Объем продукции сельского хозяйства – это макроэкономический показатель, который отражает общий объем произведенной сельскохозяйственной продукции растениеводства и животноводства за определенный период времени. Этот показатель может измеряться в натуральном (тонны, литры и т. д.) или стоимостном выражении (в денежных единицах).

Он характеризует уровень развития сельскохозяйственного сектора и его вклад в экономику страны, обеспечивая продовольственную безопасность и сырьевые ресурсы для промышленности.

Объем сельскохозяйственной продукции является аппроксимацией валовой добавленной стоимости по сельскому, лесному и рыбному хозяйству в формуле расчета ВВП по добавленной стоимости (1.3). Также из рисунка 1.1 было заключено, что доля сельскохозяйственного сектора в ВВП составляет около 7%.

3.1.4 Объем промышленного производства

Объем промышленного производства – это совокупность произведенной готовой продукции и полуфабрикатов, выполненных работ, оказанных услуг в результате осуществления видов экономической деятельности, относящихся к промышленности, за определенный период времени. Этот макроэкономический показатель может выражаться в натуральных единицах или в стоимостном выражении.

Объем промышленного производства является аппроксимацией валовой добавленной стоимости по «Горнодобывающей промышленности», «Обрабатывающей промышленности», «Снабжению электроэнергией, газом, паром, горячей водой и кондиционированным воздухом» и «Водоснабжению; сбору, обработке и удалению отходов, деятельности по ликвидации загрязнений» в формуле расчета ВВП по добавленной стоимости (1.3). Также из рисунка 1.1 было заключено, что доля промышленного сектора в ВВП составляет около 27%.

3.1.5 Объем розничного товарооборота

Объем розничного товарооборота – это макроэкономический показатель, который отражает общий объем продаж товаров конечным потребителям через розничные торговые сети, рынки и другие каналы сбыта за определенный период времени. Он измеряется в денежном выражении и включает операции, связанные с реализацией потребительских товаров населению.

Этот показатель важен для анализа потребительской активности и спроса, поскольку он характеризует уровень покупательной способности населения и тенденции в розничной торговле.

Объем розничного товарооборота является аппроксимацией компоненты C_t в формуле расчета ВВП по расходам (1.6), так как отражает покупательскую способность домашних хозяйств. Из рисунка 1.2 было заключено, что доля расходов на домашние хозяйства C_t в ВВП составляет более 50%. Также объем розничного товарооборота связан с валовой добавленной стоимостью розничной торговли из расчета ВВП по добавленной стоимости. Из рисунка 1.1 следует, что доля добавленной стоимости от оптовой и розничной торговли составляет почти 10%.

3.1.6 Объем строительно-монтажных работ

Объем строительно-монтажных работ – это макроэкономический показатель, который отражает общий объем строительных и монтажных работ, выполненных на объектах строительства за определенный период времени. Эти работы охватывают строительство новых объектов, реконструкцию, капитальный ремонт, техническое перевооружение зданий и сооружений, а также монтаж инженерных систем и оборудования.

Данный показатель выражается в стоимостном выражении и является частью общего объема инвестиций в основной капитал. Он характеризует состояние строительного сектора, так как его динамика отражает уровень инвестиционной активности и экономического развития.

Объем строительно-монтажных работ является аппроксимацией валовой добавленной стоимости по строительству в формуле расчета ВВП по добавленной стоимости (1.3). Также из рисунка 1.1 было заключено, что доля строительного сектора в ВВП составляет около 5%.

3.1.7 Сводный индекс экономических настроений

Индекс деловой активности (*Purchasing Managers' Indexes*, PMI) является часто применяемым макроэкономическим показателем для прогнозирования ВВП. Он рассчитывается на основе опроса менеджеров по закупкам. Менеджеры отвечают на вопросы, сравнивая показатель в текущем с показателем предыдущего месяца, и отвечают: «лучше», «так же» или «хуже». Сам индекс измеряется в процентах и является некоторой комбинацией переменных, соответствующих количеству менеджеров, отметивших улучшение, неизменность или сокращение деловой активности [20, 33].

Иначе говоря, PMI отражает мнение специалистов, занимающихся по-

ставкой товаров и материалов для своих компаний.

В Республике Беларусь для возможности отслеживания общеэкономической активности в целом рассчитывается *сводный индекс экономических настроений* (СИЭН), как аналог PMI. Значение этого индекса представляет собой среднее геометрическое взвешенное значение индексов экономических настроений промышленности, строительства, торговли и транспорта. Весами при взвешивании выступают доли промышленности, строительства, транспорта и торговли в ВВП. Но сам индекс, в отличие от PMI, не выражается ни в каких единицах.

Индекс СИЭН рассчитывается Национальным Банком и применяется как опережающий показатель для ВВП в силу того, что поворотные точки бизнес-цикла СИЭН опережают поворотные точки бизнес-цикла ВВП [11].

Таким образом, связь между темпами прироста ВВП и СИЭН не так очевидна, как в предыдущих случаях, поэтому мы применим эконометрические методы для исследования взаимосвязи между этими показателями. Для этого рассмотрим следующие временные ряды:

1. квартальный ВВП в средних ценах 2018 года в темпах прироста год к году,

$$iY_t = \frac{Y_t - Y_{t-4}}{Y_{t-4}} \cdot 100\%; \quad (3.1)$$

2. сезонно-сглаженный квартальный СИЭН $CESI_t$, полученный с помощью агрегации по среднему значению месячного СИЭН.

Выдвинем две гипотезы:

$$H_0 : \{\text{поведение } iY_t \text{ не зависит от } CESI_t\};$$

$$H_1 : \{\text{поведение } iY_t \text{ зависит от } CESI_t\}.$$

Рассмотрим корреляционную матрицу для рядов iY_t и $CESI_t$, представленную в таблице 3.1.

Таблица 3.1: Корреляционная матрица между $CESI_t$ и iY_t .

	$CESI_t$	iY_t
$CESI_t$	1.000000	0.779805
iY_t	0.779805	1.000000

Из таблицы 3.1 нас интересует значение коэффициента корреляции между iY_t и $CESI_t$, которое равно 0.7798. Значение коэффициента корреляции

свидетельствует в пользу того, что между временными рядами в нулевом лаге присутствует сильная линейная взаимосвязь.

Рассчитаем также кросс-корреляцию iY_t и $CESI_{t\pm k}$. В зависимости от лага k она значима до $k = 3$. То есть изменение СИЭН может сказаться на поведении темпов прироста ВВП через 3 квартала.

Построим простейшую $\text{VAR}(p)$ модель для этих двух рядов

$$\begin{cases} iY_t = \mu_1 + \sum_{j=1}^p \alpha_j \cdot iY_{t-j} + \sum_{j=1}^p \beta_j \cdot CESI_{t-j} + u_t, \\ CESI_t = \mu_2 + \sum_{j=1}^p \gamma_j \cdot iY_{t-j} + \sum_{j=1}^p \delta_j \cdot CESI_{t-j} + v_t, \end{cases} \quad (3.2)$$

где $\alpha_j, \beta_j, \gamma_j, \delta_j, \mu_1, \mu_2 \in \mathbb{R}$, $u_t, v_t \sim WN(0, \sigma^2)$. Оценив длину лагов с помощью информационных критериев AIC [21], SC [25], HQ [32] можно получить, что оптимальное число лагов $p = 1$. Таким образом, можем перестроить модель (3.2) как $\text{VAR}(1)$

$$\begin{cases} iY_t = \mu_1 + \alpha_1 \cdot iY_{t-1} + \beta_1 \cdot CESI_{t-1} + u_t, \\ CESI_t = \mu_2 + \gamma_1 \cdot iY_{t-1} + \delta_1 \cdot CESI_{t-1} + v_t. \end{cases} \quad (3.3)$$

Для построенной модели (3.3) исследуем ее свойства:

- причинность по Грейнджеру;
- функции импульсных откликов.

Результаты теста на причинность по Грейнджеру приводятся в таблице 3.2.

Таблица 3.2: Тест на причинность по Грейнджеру между iY_t и $CESI_t$.

Лаги: 1

Нулевая гипотеза:	Наблюдения	F-статистика	P-значение
$CESI_t$ не является причиной iY_t	59	5.05174	0.0286
iY_t не является причиной $CESI_t$		0.19804	0.6580

Принимая значение критического уровня равное 0.05, заключаем, что в рамках одного квартала снижение или рост СИЭН предшествуют снижению или росту темпов прироста ВВП.

Тесты на причинность на 1, 2 и 3 лагах для месячных темпов прироста ВВП и месячного СИЭН также приводят к аналогичному результату. А в силу того, что 3 лага на месячной частоте соответствуют одному лагу на квартальной частоте, получаем, что индекс СИЭН является опережающим индикатором по отношению к темпам прироста ВВП на один квартал.

Для модели (3.3) рассмотрим функции импульсов-откликов через разложение Холецкого (рисунок 3.1).

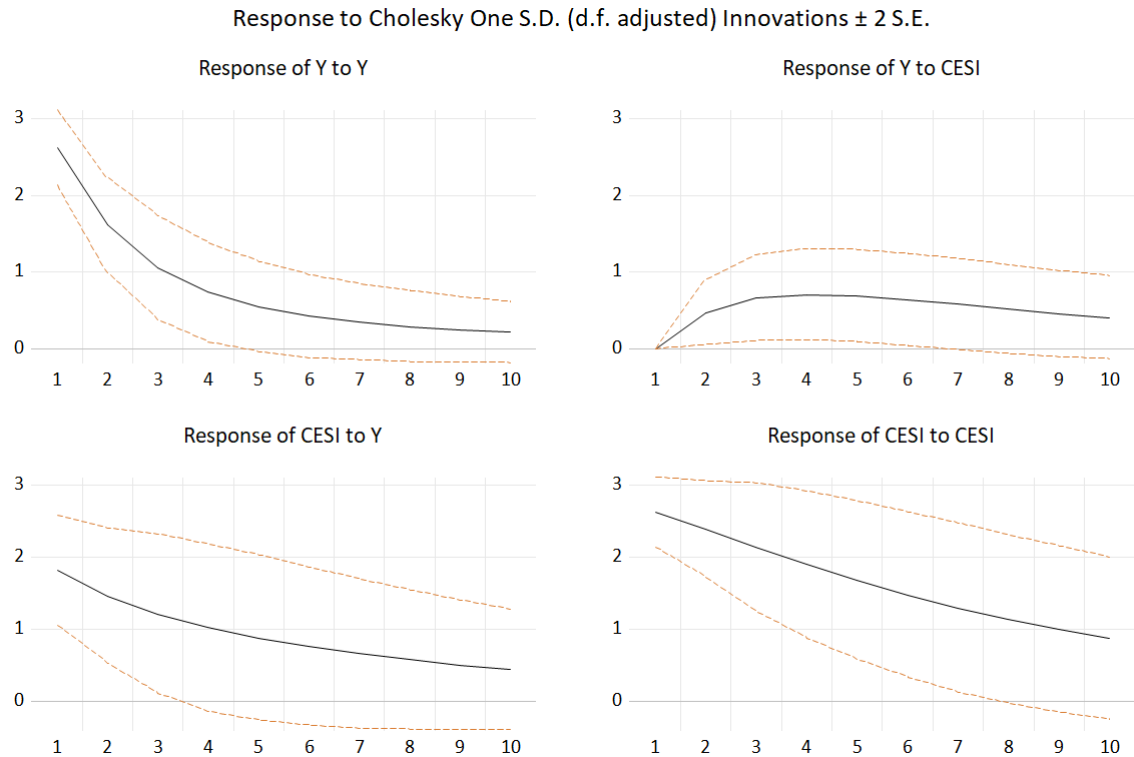


Рисунок 3.1 — Графики функций импульсов-откликов для модели (3.3)

Из рисунка 3.1 следует, что до 6-ого периода для темпов прироста ВВП \dot{Y}_t присутствуют положительные статистически значимые отклики на импульсы СИЭН.

В результате анализа мы можем заключить, что некоторая статистически значимая зависимость темпов прироста ВВП от СИЭН присутствует, следовательно, мы отклоняем гипотезу H_0 и принимаем H_1 . Таким образом, несмотря на то, что нет строгого экономического подтверждения зависимости темпов прироста ВВП от СИЭН, была доказана статистическая зависимость.

В итоге с теоретической точки зрения каждый из предоставленных месячных макроэкономических показателей связан с целевым показателем — реальным ВВП. А значит будущая модель с экономической точки зрения должна хорошо отражать реальные взаимосвязи в белорусской экономике.

3.2 Предварительный анализ временных рядов

Каждый макроэкономический показатель из параграфа 3.2 можно описать с помощью модели временного ряда. Следовательно, мы имеем 8 временных рядов:

- квартальная частота:

- $rGDP_q := Y_t^Q$ – реальный ВВП Беларуси по источникам использования доходов в среднегодовых ценах 2018 г., млн. руб.;

- месячная частота:

- $rPP_m := IP_t^M$ – объем промышленного производства в среднегодовых ценах 2018 г., млн. руб.;

- $rRet_m := RT_t^M$ – объем розничного товарооборота в среднегодовых ценах 1995 г., млн. руб.;

- $rInv_m := INV_t^M$ – объем инвестиций в основной капитал в среднегодовых ценах 2018 г., млн. руб.;

- $rAgro_m := AGRO_t^M$ – объем продукции сельского хозяйства в среднегодовых ценах 2018 г., млн. руб.;

- $Bi_Bld_m := biBLD_t^M$ – базисный индекс объема строительно-монтажных работ (янв. 2018 = 1), %;

- $Bi_rRdh_m := biINC_t^M$ – базисный индекс объема денежных доходов населения (янв. 2018 = 1). %;

- $CESI_m_SA := CESI_t^{M*}$ – сезонно-скорректированный сводный индекс экономических настроений.

Период наблюдения данных следующий

- для квартальных: I квартал 2009 года — I квартал 2025 года;
- для месячных: 1 месяц 2009 года — 1 месяц 2025 года.

Модели авторегрессии корректно работают со стационарными временными рядами. Следовательно, необходимо определить, являются ли данные временные ряды стационарными. И если они не являются стационарными, то необходимо привести их к стационарному виду, чтобы получить качественные результаты прогнозирования.

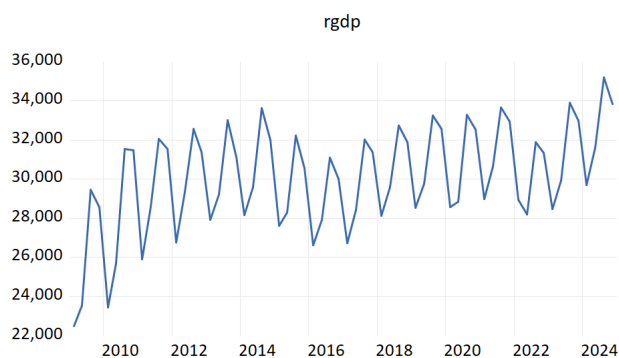


Рисунок 3.2 — График Y_t^Q

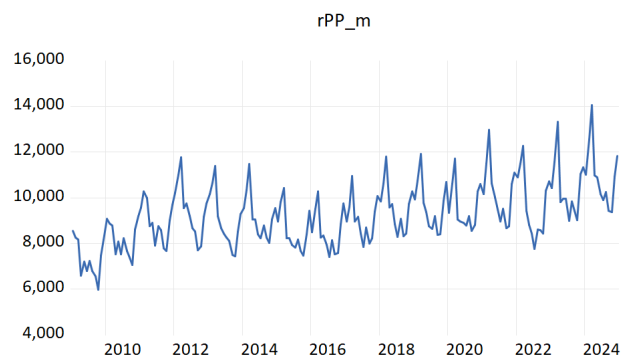


Рисунок 3.3 — График IP_t^M

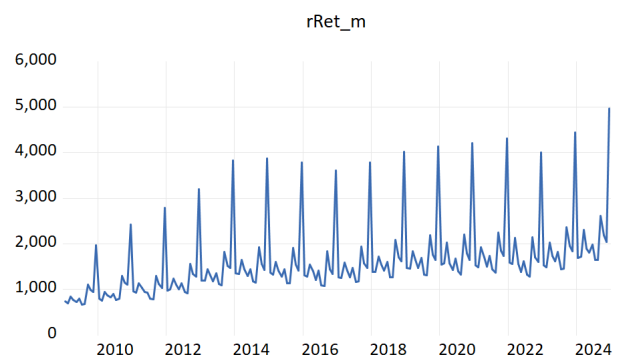


Рисунок 3.4 — График RT_t^M

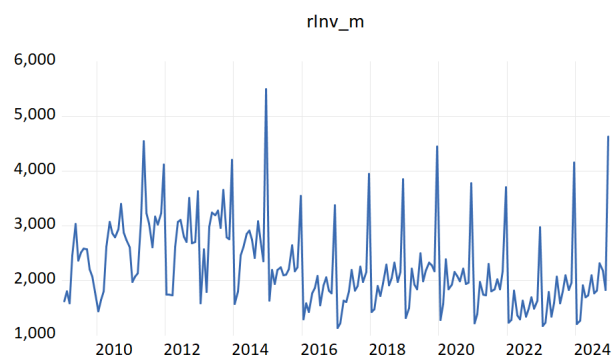


Рисунок 3.5 — График INV_t^M

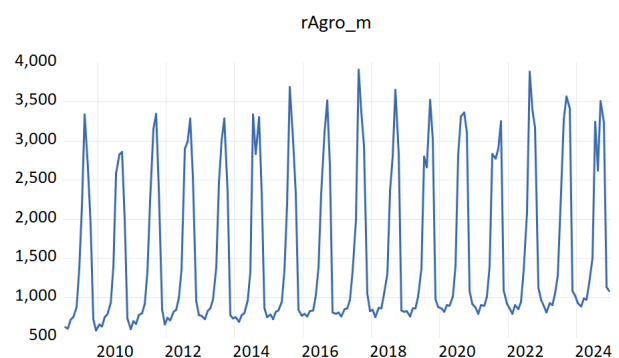


Рисунок 3.6 — График $AGRO_t^M$

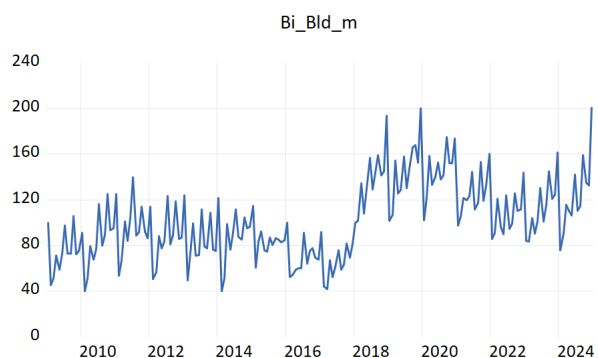


Рисунок 3.7 — График $biBLD_t^M$

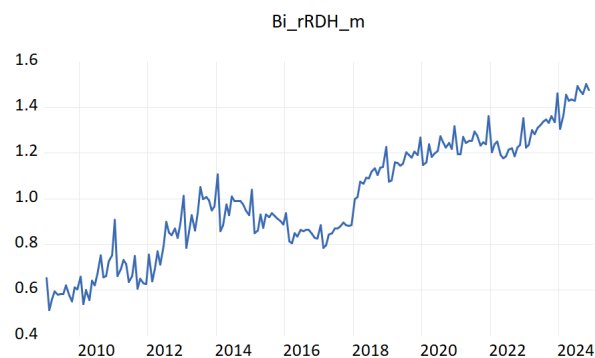


Рисунок 3.8 — График $biINC_t^M$

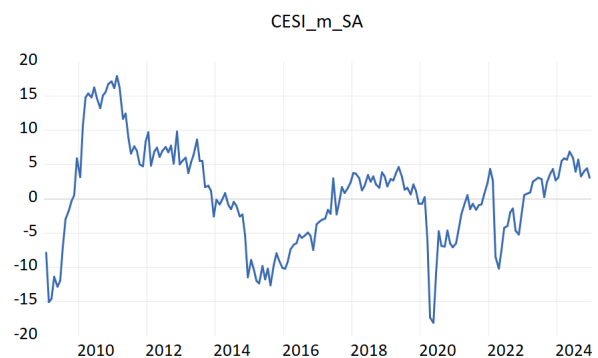


Рисунок 3.9 — График $CESI_t^{M*}$

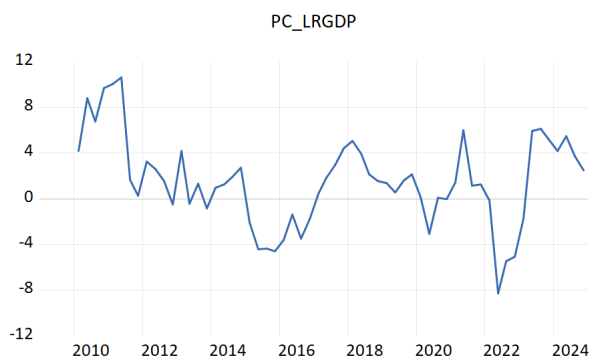


Рисунок 3.10 — График iy_t^Q

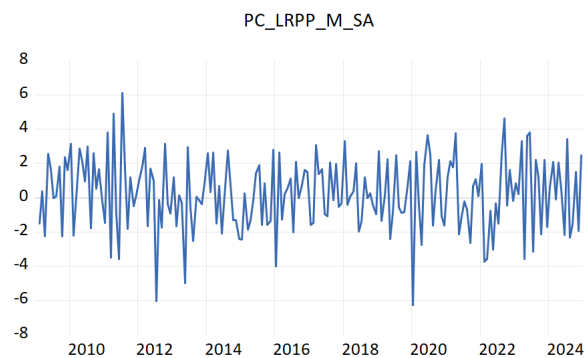


Рисунок 3.11 — График ip_t^{M*}

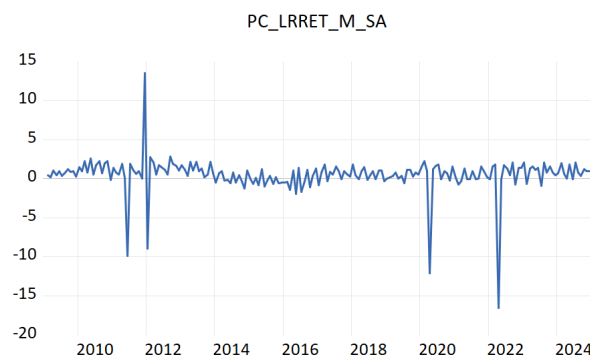


Рисунок 3.12 — График irt_t^{M*}

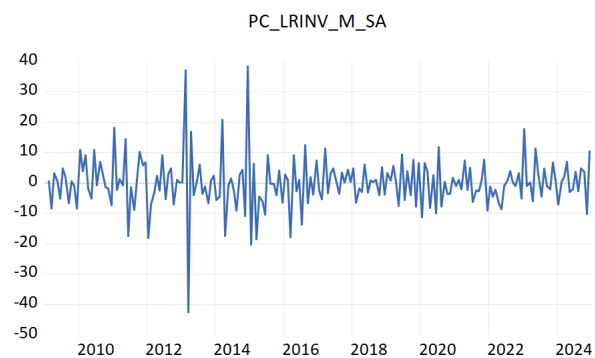


Рисунок 3.13 — График int_t^{M*}

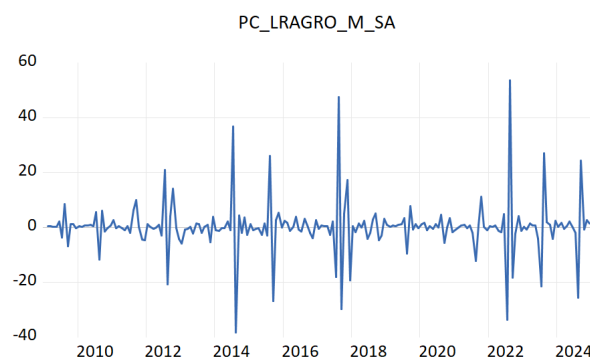


Рисунок 3.14 — График $agro_t^{M*}$

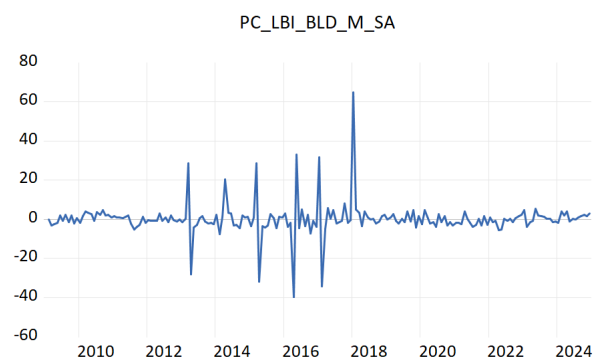


Рисунок 3.15 — График $ibld_t^{M*}$

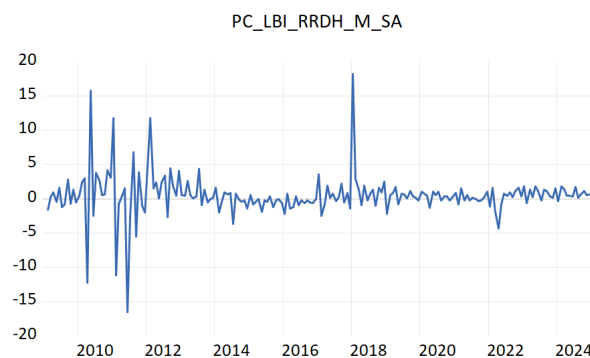


Рисунок 3.16 — График $ibinc_t^{M*}$

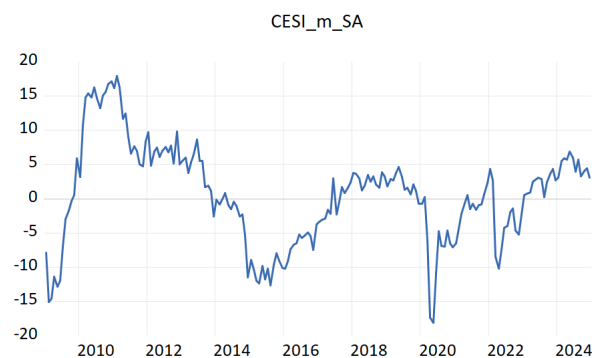


Рисунок 3.17 — График $CESI_t^{M*}$

На рисунках 3.2-3.9 приводятся графики всех исследуемых временных рядов. Из графического представления заметно, что они все кроме СИЭН обладают ярко выраженным сезонным эффектом и трендом. Следовательно, необходимо исключить из них сезонную и трендовую компоненты.

Алгоритм преобразования временных рядов следующий:

1. исключается сезонная компонента из временных рядов на месячной частоте с помощью метода TRAMO/SEATS [36] (обозначим *);
2. логарифмирование всех временных рядов (обозначим символами нижнего регистра);
3. приведение к темпам прироста путем взятия конечной разности: для целевой переменной – год к году, для месячных факторов – месяц к месяцу (обозначим i).

Таким образом, мы получили преобразованные временные ряды, графики которых представлены на рисунках 3.10-3.17. Из графического представления видно, что скачки во временных рядах резкие и мгновенные, что соответствует аддитивным выбросам (*Additive Outlier*, AO).

Для временных рядов были проведены тесты стационарности Augmented Dickey-Fuller (ADF) и Break Point Unit Root (BPUR) с аддитивным выбросом и однократным изменением в уровне [2, 12, 13, 24, 26, 39]. Тесты стационарности проверяют гипотезу H_0 : «ряд является DS-рядом (*Difference Stationary*)» против альтернативы H_1 : «ряд является TS-рядом (*Trend Stationary*)». Результаты тестирования приводятся в таблице 3.3.

Таблица 3.3: Результаты тестирования на единичный корень с помощью ADF-теста без константы и BPUR-теста с АО

Пере- менные	ADF-тест			BPUR-тест		
	t- ADF	p- значение	H_0 : DS-ряд	t- ADF	p- значение	H_0 : DS-ряд
i y_t^Q	-2.80	0.006	отвергается	-4.49	0.045	отвергается
i ip_t^{M*}	-17.24	0.000	отвергается	-18.21	< 0.01	отвергается
i rt_t^{M*}	-15.52	0.000	отвергается	-19.11	< 0.01	отвергается
i inv_t^{M*}	-23.56	0.000	отвергается	-24.33	< 0.01	отвергается
i $agro_t^{M*}$	-11.35	0.000	отвергается	-25.94	< 0.01	отвергается
i $bi bld_t^{M*}$	-17.64	0.000	отвергается	-18.97	< 0.01	отвергается
i $bi inc_t^{M*}$	-16.23	0.000	отвергается	-18.13	< 0.01	отвергается
$CESI_t^{M*}$	-2.52	0.012	отвергается	-3.42	0.429	не отвергается

По результатам тестирования из таблицы 3.3, несмотря на то, что BPUR-тест позволяет принять гипотезу о нестационарности, мы отклоняем гипотезу в силу экономического смысла этого ряда. Тогда каждый из временных рядов является TS-рядом. Следовательно, по ним можно строить адекватные модели авторегрессии.

3.3 Теоретическое описание макроэкономической модели ВВП

В таблице 3.4 приведены все переменные, которые будут участвовать в построении математической модели

Таблица 3.4: Переменные модели

Переменная	Обработка	Описание	Календарь вы- хода
$i y_t^Q$	Сезонная коррек- тировка, темпы прироста год к го- ду в логарифмах	Реальный ВВП Беларуси по ис- точникам исполь- зования доходов в среднегодовых ценах 2018 г.	Задержка 90 дней после квартально- го периода + кор- ректировка в де- кабре года, следу- ющего за отчет- ным
$i i p_t^{M*}$	Сезонная коррек- тировка, темпы прироста меся- ца к месяцу в логарифмах	Объем про- мышленного производства в среднегодовых ценах, 2018 г.	Задержка 17 дней после месячного периода
$i r t_t^{M*}$	Сезонная коррек- тировка, темпы прироста меся- ца к месяцу в логарифмах	Объем рознично- го товарооборота в среднегодовых ценах 1995 г.	Задержка 18 дней после месячного периода
$i i n v_t^{M*}$	Сезонная коррек- тировка, темпы прироста меся- ца к месяцу в логарифмах	Объем инвести- ций в основной капитал в сред- негодовых ценах 2018 г.	Задержка 24 дня после месячного периода

$i agro_t^{M*}$	Сезонная корректировка, темпы прироста месяц к месяцу в логарифмах	Объем продукции сельского хозяйства в среднегодовых ценах 2018 г.	Задержка 19 дней после месячного периода
$i bi bld_t^{M*}$	Сезонная корректировка, темпы прироста месяц к месяцу в логарифмах	Базисный индекс объема строительно монтажных работ (январь 2018 = 1)	Задержка 21 день после месячного периода
$i bi inc_t^{M*}$	Сезонная корректировка, темпы прироста месяц к месяцу в логарифмах	Базисный индекс объема денежных доходов населения (январь 2018 = 1)	Задержка 30 дней после месячного периода
$CESI_t^{M*}$	Сезонная корректировка	Сводный индекс экономических настроений	Задержка 4 дня после месячного периода

Всего в модели ВВП имеется 8 переменных:

- одна эндогенная переменная на квартальной частоте;
- семь экзогенных переменных на месячной частоте.

Теоретически модель ВВП можно представить в виде схемы, представленной на рисунке 3.18.

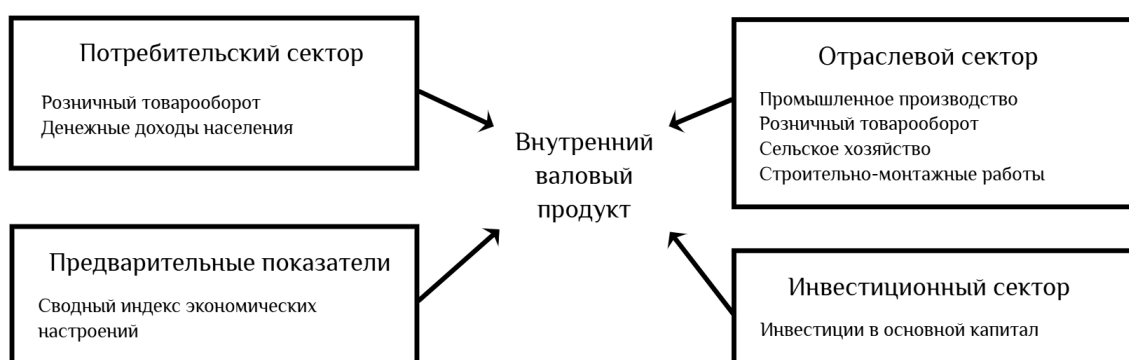


Рисунок 3.18 — Схема теоретических связей в модели ВВП

То есть модель ВВП строится по трем секторам экономики Беларуси, а также по одному предварительному показателю.

Строить математическую модель ВВП мы будем на основе MF-VAR модели. В параграфе 2.2 мы заключили:

- модель MF-VAR представляет каждую переменную на месячной частоте в виде трех переменных на квартальной частоте;
- все переменные в модели векторной авторегрессии являются эндогенными.

Следовательно, каждая из семи экзогенных переменных на месячной частоте раскладывается в три переменные на квартальной частоте, и в итоге мы получаем 21 эндогенную переменную на квартальной частоте. Также все регрессионные модели включают константу c , которая соответствует отклонению значений модели от нуля.

Также обратим внимание на график темпов прироста ВВП, представленный на рисунке 3.10. Во II-ом квартале 2022-ого года во временном ряде присутствует аномальный скачок. В силу того, что это мгновенный структурный сдвиг, модель не сможет адекватно спрогнозировать такое поведение временного ряда. Поэтому в модель включается также импульсная фиктивная переменная (*dummy variable*), которая будет «подсказывать» модели, что в этой точке происходит структурный сдвиг, что позволит не влиять этому скачку на результаты прогноза. Импульсная экзогенная переменная dum2022q2 задана на квартальной частоте таким образом, что в точке 2022Q2 она принимает значение «1», а во всех остальных точках – «0».

Таким образом, в силу описанных условий, модель ВВП на основе MF-VAR будет включать 24 переменные:

- двадцать две эндогенные переменные на квартальной частоте;
- две экзогенные переменные на квартальной частоте.

Следовательно, в общем случае модель представляется в виде системы из 22 уравнений, соответствующих каждой из эндогенных переменных.

Эконометрическая модель ВВП на основе MF-VAR будет задаваться в следующем виде

$$\mathbf{X}_t = \mathbf{c} + \mathbf{c}_1 \times \text{dum2022q2} + \sum_{k=1}^p \mathbf{B}_k \times \mathbf{X}_{t-k} + \boldsymbol{\varepsilon}_t, \quad (3.4)$$

где $\mathbf{X}_t = \begin{bmatrix} i y_t^Q, i p_t^{M_1*}, i p_t^{M_2*}, i p_t^{M_3*}, i r_t^{M_1*}, i r_t^{M_2*}, i r_t^{M_3*}, i n v_t^{M_1*}, i n v_t^{M_2*}, i n v_t^{M_3*}, i a g r o_t^{M_1*}, i a g r o_t^{M_2*}, i a g r o_t^{M_3*}, i b i b l d_t^{M_1*}, i b i b l d_t^{M_2*}, i b i b l d_t^{M_3*}, i b i i n c_t^{M_1*}, i b i i n c_t^{M_2*}, i b i i n c_t^{M_3*}, C E S I_t^{M_1*}, C E S I_t^{M_2*}, C E S I_t^{M_3*} \end{bmatrix}$; \mathbf{c} , \mathbf{c}_1 – векторные параметры модели, $\mathbf{B}_k = (\beta_k^{i,j})$ – матрица параметров модели раз-

мерности 22×22 , ε_t – вектор внешних шоков из $WN(0, \sigma^2)$, p – количество лагов модели. В \mathbf{X}_t все временные ряды заданы на квартальной частоте.

В итоге для построения модели необходимо оценить $22 \times (2 + 22 \times p)$ коэффициентов, поскольку производится оценка параметров \mathbf{c} , \mathbf{c}_1 и каждой матрицы параметров \mathbf{B}_k , $k = 1, \dots, p$. Оценка коэффициентов производится с помощью метода наименьших квадратов при заданном периоде оценивания T .

Параметризация и оценка прогнозов модели производятся с помощью алгоритма расширяющегося окна, описанного в подпараграфе 2.3.3.

Период для оценки коэффициентов был выбран с 2009Q1 до 2022Q2, то есть в начальной выборке присутствует 54 наблюдения. Это минимальное число наблюдений для выборки, так как при объеме выборки меньше 54 матрица коэффициентов \mathbf{B}_k оказывается сингулярной, следовательно коэффициенты модели невозможно оценить по методу наименьших квадратов.

Вся выборка состоит из 64 наблюдений, поскольку наблюдаемый период взят от 2009Q1 до 2024Q4. Следовательно, если для первоначальной оценки параметров отводится 54 наблюдения, то в алгоритме расширяющегося окна будут использоваться оставшиеся 10 наблюдений от 2022Q3 до 2024Q4.

Таким образом, с помощью модели ВВП будут построены краткосрочные прогнозы на 10 кварталов вперед. На этих 10 точках будут рассчитаны основные метрики оценки качества моделей, что даст возможность провести сравнительный анализ моделей с разным набором параметров: при разном подборе переменных и лагов.

3.4 Результаты построения и использования моделей в задачах экономического анализа и прогнозирования

3.4.1 Формальное описание моделей и результаты оценивания коэффициентов

В процессе исследования для сравнительного анализа были построены две модели MF-VAR, а также две модели VAR с агрегированием. Построенные модели имеют следующие спецификации

- VAR₁(1) – это VAR модель с 1 лагом по агрегированным переменным $ip_t^{M*}, irt_t^{M*}, inv_t^{M*}, agro_t^{M*}, ibibld_t^{M*}, ibiinc_t^{M*}, CESI_t^{M*}$;

- VAR₂(2) – это VAR модель с 2 лагами по агрегированной переменной ip_t^{M*} ; она обладает наилучшими значениями метрик RMSE, MAE на вневыборочных прогнозах в классе VAR моделей;
- MF-VAR₁(1) – это MF-VAR модель с 1 лагом по переменным ip_t^{M*} , irt_t^{M*} , inv_t^{M*} , $agro_t^{M*}$, $ibld_t^{M*}$, $ibinc_t^{M*}$, $CESI_t^{M*}$;
- MF-VAR₂(2) – это MF-VAR модель с 2 лагами по переменным ip_t^{M*} , irt_t^{M*} , $CESI_t^{M*}$; она обладает наилучшими значениями метрик RMSE, MAE на вневыборочных прогнозах в классе MF-VAR моделей.

Для наилучшей модели MF-VAR₂(2) опишем ее формальную структуру. Эта модель задается в виде системы уравнений

$$\begin{aligned} iy_t^Q = & c^1 + c_1^1 \text{dum2022q2}_t + \sum_{k=1}^2 \sum_{j=1}^3 \alpha_k^{1,j} ip_{t-k}^{M_{j*}} + \\ & + \sum_{k=1}^2 \sum_{j=1}^3 \beta_k^{1,j} irt_{t-k}^{M_{j*}} + \sum_{k=1}^2 \sum_{j=1}^3 \gamma_k^{1,j} CESI_{t-k}^{M_{j*}} + \sum_{k=1}^2 \delta_k^1 iy_{t-k}^Q, \end{aligned}$$

$$\begin{aligned} irt_t^{M_{1*}} = & c^2 + c_1^2 \text{dum2022q2}_t + \sum_{k=1}^2 \sum_{j=1}^3 \alpha_k^{2,j} ip_{t-k}^{M_{j*}} + \\ & + \sum_{k=1}^2 \sum_{j=1}^3 \beta_k^{2,j} irt_{t-k}^{M_{j*}} + \sum_{k=1}^2 \sum_{j=1}^3 \gamma_k^{2,j} CESI_{t-k}^{M_{j*}} + \sum_{k=1}^2 \delta_k^2 iy_{t-k}^Q, \end{aligned}$$

$$\begin{aligned} irt_t^{M_{2*}} = & c^3 + c_1^3 \text{dum2022q2}_t + \sum_{k=1}^2 \sum_{j=1}^3 \alpha_k^{3,j} ip_{t-k}^{M_{j*}} + \\ & + \sum_{k=1}^2 \sum_{j=1}^3 \beta_k^{3,j} irt_{t-k}^{M_{j*}} + \sum_{k=1}^2 \sum_{j=1}^3 \gamma_k^{3,j} CESI_{t-k}^{M_{j*}} + \sum_{k=1}^2 \delta_k^3 iy_{t-k}^Q, \end{aligned}$$

$$\begin{aligned} irt_t^{M_{3*}} = & c^4 + c_1^4 \text{dum2022q2}_t + \sum_{k=1}^2 \sum_{j=1}^3 \alpha_k^{4,j} ip_{t-k}^{M_{j*}} + \\ & + \sum_{k=1}^2 \sum_{j=1}^3 \beta_k^{4,j} irt_{t-k}^{M_{j*}} + \sum_{k=1}^2 \sum_{j=1}^3 \gamma_k^{4,j} CESI_{t-k}^{M_{j*}} + \sum_{k=1}^2 \delta_k^4 iy_{t-k}^Q, \end{aligned}$$

$$\begin{aligned} \mathbf{i} ip_t^{M_1*} &= c^5 + c_1^5 \text{dum2022q}2_t + \sum_{k=1}^2 \sum_{j=1}^3 \alpha_k^{5,j} \mathbf{i} ip_{t-k}^{M_j*} + \\ &+ \sum_{k=1}^2 \sum_{j=1}^3 \beta_k^{5,j} \mathbf{i} rt_{t-k}^{M_j*} + \sum_{k=1}^2 \sum_{j=1}^3 \gamma_k^{5,j} CESI_{t-k}^{M_j*} + \sum_{k=1}^2 \delta_k^5 \mathbf{i} y_{t-k}^Q, \end{aligned}$$

$$\begin{aligned} \mathbf{i} ip_t^{M_2*} &= c^6 + c_1^6 \text{dum2022q}2_t + \sum_{k=1}^2 \sum_{j=1}^3 \alpha_k^{6,j} \mathbf{i} ip_{t-k}^{M_j*} + \\ &+ \sum_{k=1}^2 \sum_{j=1}^3 \beta_k^{6,j} \mathbf{i} rt_{t-k}^{M_j*} + \sum_{k=1}^2 \sum_{j=1}^3 \gamma_k^{6,j} CESI_{t-k}^{M_j*} + \sum_{k=1}^2 \delta_k^6 \mathbf{i} y_{t-k}^Q, \end{aligned}$$

$$\begin{aligned} \mathbf{i} ip_t^{M_3*} &= c^7 + c_1^7 \text{dum2022q}2_t + \sum_{k=1}^2 \sum_{j=1}^3 \alpha_k^{7,j} \mathbf{i} ip_{t-k}^{M_j*} + \\ &+ \sum_{k=1}^2 \sum_{j=1}^3 \beta_k^{7,j} \mathbf{i} rt_{t-k}^{M_j*} + \sum_{k=1}^2 \sum_{j=1}^3 \gamma_k^{7,j} CESI_{t-k}^{M_j*} + \sum_{k=1}^2 \delta_k^7 \mathbf{i} y_{t-k}^Q, \end{aligned}$$

$$\begin{aligned} CESI_t^{M_1*} &= c^8 + c_1^8 \text{dum2022q}2_t + \sum_{k=1}^2 \sum_{j=1}^3 \alpha_k^{8,j} \mathbf{i} ip_{t-k}^{M_j*} + \\ &+ \sum_{k=1}^2 \sum_{j=1}^3 \beta_k^{8,j} \mathbf{i} rt_{t-k}^{M_j*} + \sum_{k=1}^2 \sum_{j=1}^3 \gamma_k^{8,j} CESI_{t-k}^{M_j*} + \sum_{k=1}^2 \delta_k^8 \mathbf{i} y_{t-k}^Q, \end{aligned}$$

$$\begin{aligned} CESI_t^{M_2*} &= c^9 + c_1^9 \text{dum2022q}2_t + \sum_{k=1}^2 \sum_{j=1}^3 \alpha_k^{9,j} \mathbf{i} ip_{t-k}^{M_j*} + \\ &+ \sum_{k=1}^2 \sum_{j=1}^3 \beta_k^{9,j} \mathbf{i} rt_{t-k}^{M_j*} + \sum_{k=1}^2 \sum_{j=1}^3 \gamma_k^{9,j} CESI_{t-k}^{M_j*} + \sum_{k=1}^2 \delta_k^9 \mathbf{i} y_{t-k}^Q, \end{aligned}$$

$$CESI_t^{M_3^*} = c^{10} + c_1^{10} \text{dum2022q2}_t + \sum_{k=1}^2 \sum_{j=1}^3 \alpha_k^{10,j} i p_{t-k}^{M_j^*} + \\ + \sum_{k=1}^2 \sum_{j=1}^3 \beta_k^{10,j} i r t_{t-k}^{M_j^*} + \sum_{k=1}^2 \sum_{j=1}^3 \gamma_k^{10,j} CESI_{t-k}^{M_j^*} + \sum_{k=1}^2 \delta_k^{10} i y_{t-k}^Q,$$

где $c^i, c_1^i, \alpha_k^{i,j}, \beta_k^{i,j}, \gamma_k^{i,j}, \delta_k^i \in \mathbb{R}$, $i = \overline{1, 10}$ – коэффициенты, определяемые по методу наименьших квадратов.

Из формальной записи модели видно, что требуется оценить 220 коэффициентов и проверить адекватности 10 уравнений. Но так как нас в первую очередь интересует именно описание прироста ВВП, то для оценки адекватности модели мы будем ориентироваться в первую очередь на уравнение для ВВП.

После оценки коэффициентов для уравнения ВВП мы получаем следующие статистики, приведенные в таблице 3.5.

Таблица 3.5: Статистики по модели MF-VAR₂(2)

Показатель	Значение
По уравнению для $i y_t^Q$	
Коэффициент детерминации (R-squared)	0.88
Скорректированный R-squared	0.80
F-статистика	12.02
По всей модели	
Определитель ковариации остатков (с учетом степени свободы)	3.39E-21
Определитель ковариации остатков	2.88E-23
Критерий Акаике (AIC)	-15.93
Критерий Шварца (SC)	-8.12
Число коэффициентов	220

Из таблицы 3.5 мы можем сделать следующие выводы:

- коэффициент детерминации для уравнения прироста ВВП достаточно высок, значит, модель достаточно хорошо объясняет дисперсию прироста ВВП;
- значение F-статистики достаточно высокое, что свидетельствует о значимости уравнения внутри системы в целом;
- значение детерминанта ковариационной матрицы остатков достаточно низкое, что говорит о том, что модель адекватно описывает зависимости между переменными;

- значения AIC и SC для модели в целом достаточно низкие, что указывает на то, что модель достаточно качественная.

Также для оценки адекватности модели необходимо провести анализ остатков. Для оцененной модели в таблице 3.6 представлен результат тестирования гипотезы о законе распределения остатков.

Таблица 3.6: Результаты тестирования гипотезы подчинении остатков нормальному закону распределения $\mathcal{N}(0, \sigma^2)$

Метод	Значение статистики	P-значение	Решение
Cramer-von Mises (W^2)	0.0442	> 0.25	принимается
Watson (U^2)	0.0437	> 0.25	принимается
Anderson-Darling (A^2)	0.315125	> 0.25	принимается

Таким образом, по результатам тестирования остатков на подчинение нормальному закону распределения с нулевым средним, приведенным в таблице 3.6, мы заключаем, что остатки подчиняются нормальному закону.

Для остатков также был проведен тест на автокорреляцию. По итогам теста автокорреляция в остатках отсутствует. Следовательно, мы можем заключить, что поведение остатков соответствует корректно построенной модели.

А значит модель является статистически значимой и ее можно использовать для прогнозирования и наукастинга по смешанным данным.

3.4.2 Сравнительный анализ моделей по результатам прогнозов

Для проведения сравнительного анализа было разработано программное обеспечение в среде Eviews для реализации алгоритма оценки точности по методу расширяющегося окна. Для каждой из моделей был построен ретроспективный прогноз с помощью метода расширяющегося окна: модель была оценена на всем временном периоде, а затем последовательно строились краткосрочные прогнозы. После чего с помощью этого же метода для каждой из моделей были построены вневыборочные прогнозы.

На рисунках 3.19-3.26 представлены графики ретроспективных прогнозов (в левом столбце) и вневыборочных прогнозов (в правом столбце) соответствующих моделей по методу расширяющегося окна.

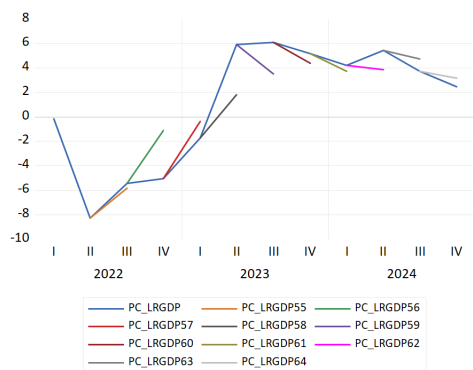


Рисунок 3.19 — Ретроспективный прогноз VAR₁(1)

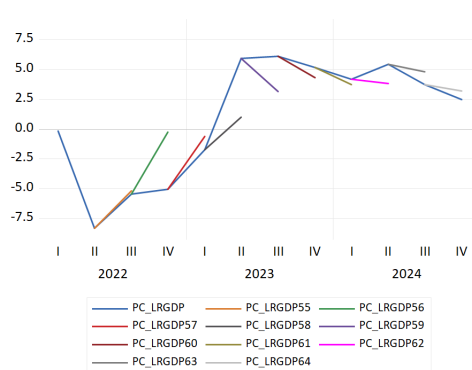


Рисунок 3.20 — Вневыборочный прогноз VAR₁(1)

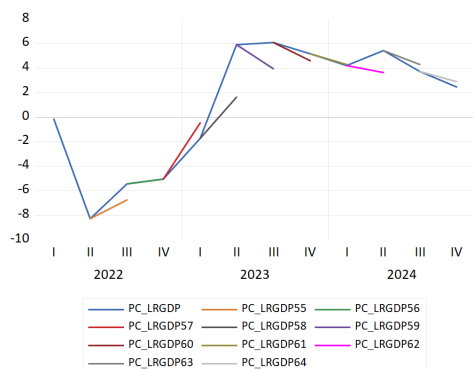


Рисунок 3.21 — Ретроспективный прогноз VAR₂(2)

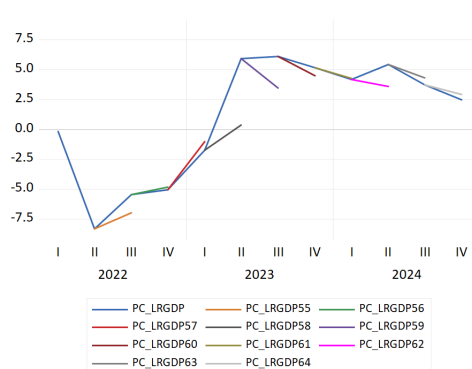


Рисунок 3.22 — Вневыборочный прогноз VAR₂(2)

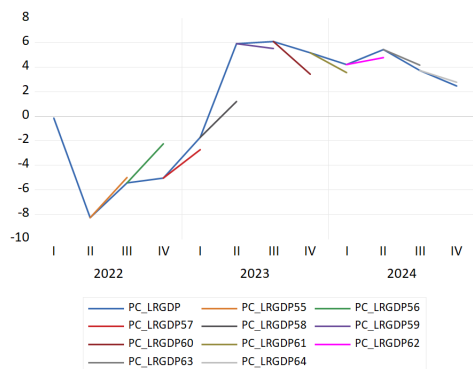


Рисунок 3.23 — Ретроспективный прогноз MF-VAR₁(1)

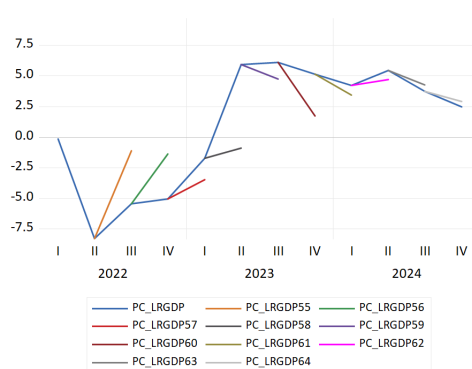


Рисунок 3.24 — Вневыборочный прогноз MF-VAR₁(1)

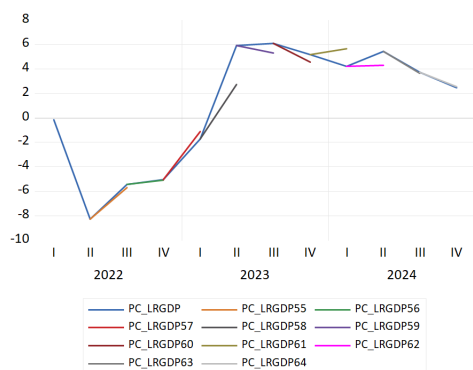


Рисунок 3.25 — Ретроспективный прогноз MF-VAR₂(2)

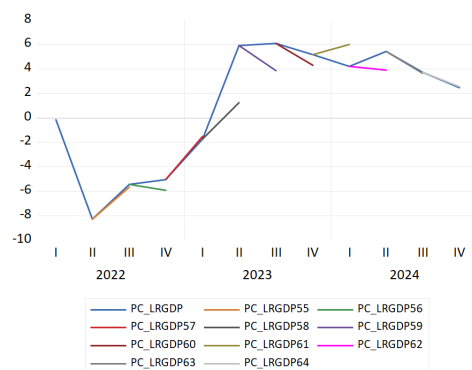


Рисунок 3.26 — Вневыборочный прогноз MF-VAR₂(2)

Из графиков на рисунках 3.19-3.26 можно заключить:

- на более нестабильных периодах (2022-2023 гг.) как ретроспективные, так и вневыборочные прогнозы становятся хуже;
- на более стабильных периодах (2024 г.) как ретроспективные, так и вневыборочные прогнозы оказываются достаточно точными;
- наилучшая модель MF-VAR₂(2) по своим вневыборочным прогнозам оказывается достаточно близкой к ретроспективным прогнозам, что также подкрепляет наши наблюдения об устойчивости модели.

Для численной оценки результатов прогнозов рассчитаны метрики RMSE и MAE. Все значения метрик, полученные к ходе экспериментов, приведены в таблице 3.7.

Таблица 3.7: Показатели точности прогнозов годовых темпов прироста ВВП Беларуси

Модель	RMSE	MAE
Прогнозный период 2022Q3-2024Q4 (ретроспективные прогнозы)		
VAR ₁ (1)	2.1391	1.6827
VAR ₂ (2)	1.7437	1.2460
MF-VAR ₁ (1)	1.8983	1.3382
MF-VAR ₂ (2)	1.2306*	0.8253*
Прогнозный период 2022Q3-2024Q4 (вневыборочные прогнозы)		
VAR ₁ (1)	2.4954	1.8776
VAR ₂ (2)	2.1128	1.4246
MF-VAR ₁ (1)	3.1069	2.3759
MF-VAR ₂ (2)	1.8467*	1.2564*

Во-первых, сравнивая значения метрик на ретроспективных и вневыборочных прогнозах, заметим, что значения метрик на вневыборочных прогнозах не слишком сильно отклоняются от ретроспективных, что свидетельствует об устойчивости построенных моделей.

Во-вторых, по итогам сравнительного анализа построенных моделей было заключено, что модель MF-VAR₂(2) обладает наилучшими прогностическими способностями в классе рассматриваемых векторных авторегрессионных моделей в смысле RMSE и MAE метрик как на ретроспективных, так и на вневыборочных прогнозах.

Таким образом, модель MF-VAR₂(2) является наилучшей в классе моделей векторной авторегрессии по данному набору макроэкономических показателей.

3.4.3 Характер взаимосвязей в модели MF-VAR₂(2)

Характер взаимосвязей в модели мы будем оценивать с помощью анализа функций импульсных откликов и причинности по Грейнджеру.

Для подкрепления предположения о значимых взаимосвязях в модели был проведен тест на причинность по Грейнджеру на всей модели. Однако нам в первую очередь интересно, могут ли высокочастотные показатели объяснять прирост ВВП. Результаты теста по приросту ВВП приведены в таблице 3.8.

Таблица 3.8: Тест на причинность PC_LR GDP в зависимости от остальных показателей для модели MF-VAR₂(2)

Показатель	χ^2 - статистика	Р-значение	Решение
PC_LRPP_M_SA_1	10.28057	0.0059	является причиной
PC_LRPP_M_SA_2	9.796988	0.0075	является причиной
PC_LRPP_M_SA_3	7.762518	0.0206	является причиной
PC_LRRET_M_SA_1	12.94955	0.0015	является причиной
PC_LRRET_M_SA_2	7.665317	0.0217	является причиной
PC_LRRET_M_SA_3	7.001314	0.0302	является причиной
CESI_M_SA_1	8.863230	0.0119	является причиной
CESI_M_SA_2	9.342520	0.0094	является причиной
CESI_M_SA_3	2.445332	0.2944	не является причиной

По результатам тестирования на причинность по Грейнджеру, приведенным в таблице 3.8, можно заключить, что все все показатели, кроме 3-его месяца СИЭН являются причиной по Грейнджеру для прироста ВВП. Несмотря на это 1-ый и 2-ой месяцы СИЭН объясняют прирост ВВП в квартале, поэтому мы заключаем, что все 3 высокочастотные переменные объясняют прирост ВВП в текущем квартале.

На рисунке 3.27 представлены графики функций откликов ВВП на импульсы переменных в модели MF-VAR₂(2).

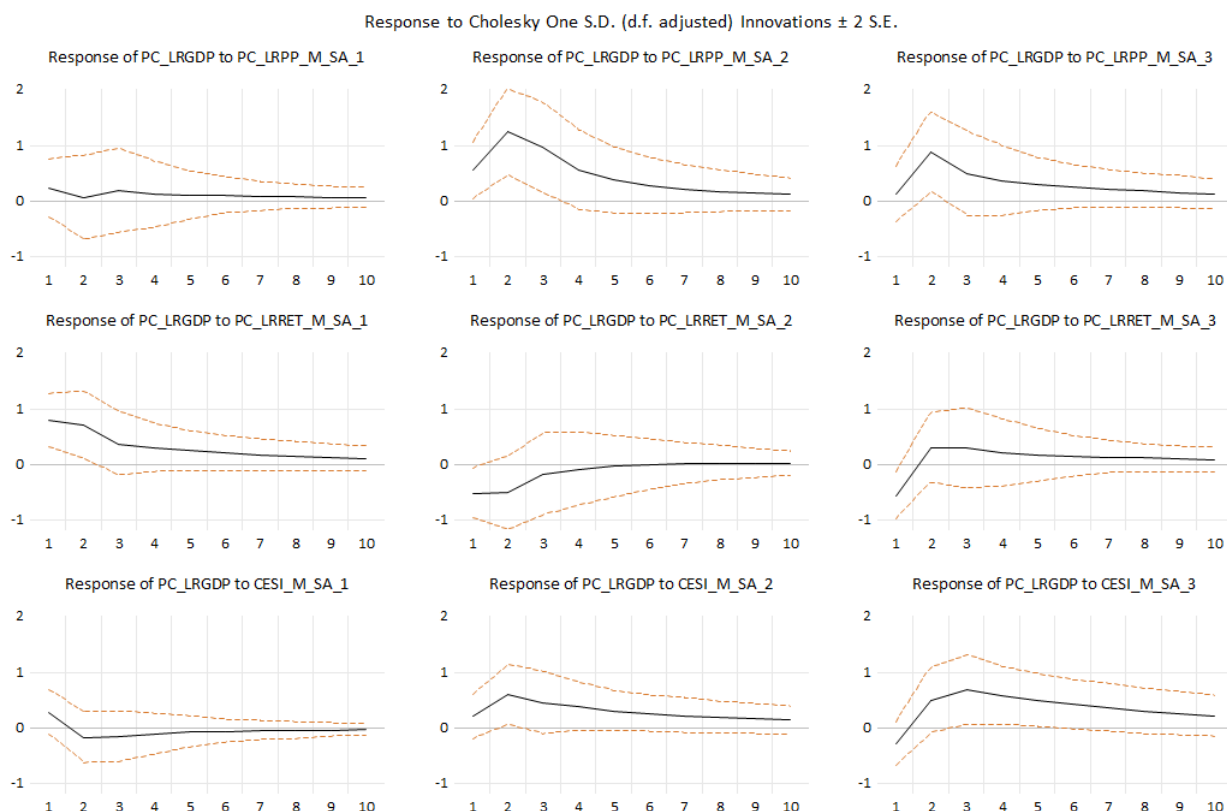


Рисунок 3.27 — Графики откликов прироста ВВП на импульсы каждого показателя для модели ВВП на основе MFVAR₂(2)

Из графиков функций импульсных откликов следует, что ВВП имеет положительные статистически значимые отклики на импульсы каждого из включенных в эту модель показателей. Это свойство позволяет установить соответствие с экономическим смыслом модели.

Из всех построенных графиков импульсных откликов также заметно, что отклики значимы на относительно коротких промежутках: не больше 1-2 периодов. Следовательно, каждый показатель оказывает лишь краткосрочный эффект на темпы прироста ВВП. Это подкрепляет теоретические предположения о том, что построенная модель MF-VAR показывает себя лучше именно в задачах краткосрочного прогнозирования.

В соответствии с полученными результатами из импульсных откликов и причинности по Грейнджеру делаем вывод, что все показатели в модели MF-VAR₂(2) действительно имеют опережающий характер для прироста ВВП. Это также свидетельствует в пользу адекватности модели.

3.4.4 Построение наукастинга

В качестве итогового эксперимента был произведен наукастинг при помощи модели MF-VAR₂(2). Алгоритм эксперимента был следующим

- во всех временных рядах месячной частоты каждый 2-ой и 3-ий месяц квартала были заменены на 1-ый месяц квартала; с помощью расширяющегося окна были построены прогнозы;
- в каждый временной ряд был возвращен 2-ой месяц квартала, а 3-ий месяц был заменен на 2-ой месяц квартала; снова строились прогнозы;
- в каждый временной ряд был возвращен 3-ий месяц квартала; снова строились прогнозы.

Результаты прогнозов с помощью расширяющегося окна приведены на рисунке 3.28.

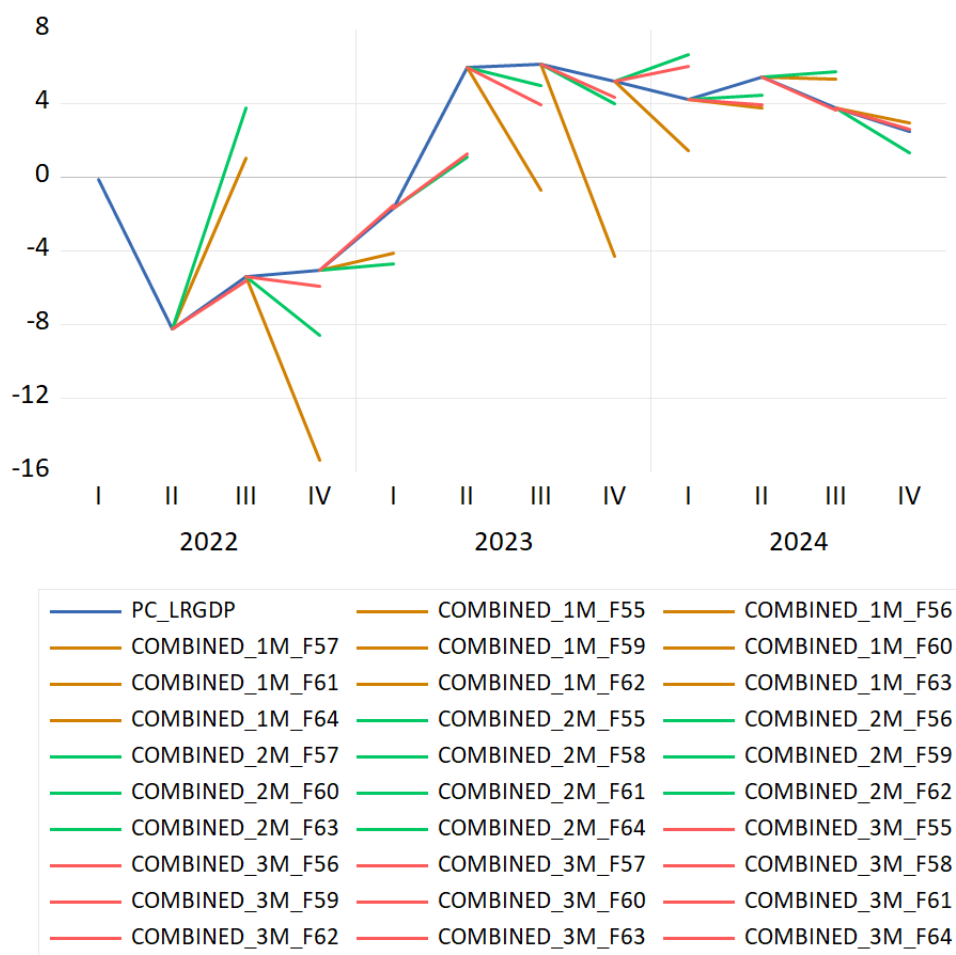


Рисунок 3.28 — Вневыборочный наукастинг для модели MF-VAR₂(2)

В таблице 3.9 представлены значения метрик в каждой точке наукаста.

Таблица 3.9: Таблица сравнения RMSE и MAE по кварталам

Дата	RMSE			MAE		
	1 месяц	2 месяца	3 месяца	1 месяц	2 месяца	3 месяца
2022Q3	41.73699	84.39614	0.04388	6.460418	9.186737	0.209475
2022Q4	106.8192	12.6813	0.789954	10.33534	3.561081	0.888794
2023Q1	5.882098	8.950058	0.041751	2.425304	2.991665	0.204331
2023Q2	27.44998	23.39465	21.96949	5.239273	4.836802	4.687163
2023Q3	47.03265	1.328504	4.930977	6.858035	1.152607	2.220582
2023Q4	90.0033	1.448812	0.751676	9.487007	1.203666	0.866993
2024Q1	7.898813	5.846431	3.230362	2.810428	2.417939	1.797321
2024Q2	2.846819	1.075519	2.332993	1.687252	1.037027	1.527414
2024Q3	2.510508	9.02388	0.006636	1.584316	1.97556	0.081463
2024Q4	0.204018	1.448143	0.006425	0.451684	1.203388	0.080155

Из графика и таблицы можно заключить, что модель более устойчива в наукастинге на спокойных периодах, как, например, в 2024-ом году. На этом периоде прогнозы близки друг к другу. На нестабильных периодах, как в 2022-ом году, прогнозы по 1-ому и 2-ому месяцам очень сильно отклоняются как от прогноза по 3-м месяцам, так и от реального значения временного ряда.

Таким образом, модель MF-VAR₂(2) по опережающим показателям «промышленное производство», «розничный товарооборот», «СИЭН» обладает наилучшими прогностическими способностями среди моделей авторегрессии, соответствует экономическому смыслу и является адекватной в смысле остатков модели. Она также пригодна для наукастинга в ситуациях, когда экономика Беларуси ведет себя стабильно, без резких скачков или падений.

ЗАКЛЮЧЕНИЕ

В дипломной работе получены следующие основные результаты:

- 1) сформулирован и подготовлен обзор основных подходов и методов, применяемых для краткосрочного прогнозирования и наукастинга;
- 2) подготовлено описание эконометрических моделей и методов, применяемых для прогнозирования прироста ВВП Республики Беларусь;
- 3) разработан алгоритмический инструментарий и программное обеспечение на основе метода расширяющегося окна для краткосрочного прогнозирования по моделям векторной авторегрессии по смешанным данным;
- 4) проведен сравнительный анализ моделей по смешанным данным с моделями по агрегированным данным.

На основе полученных результатов можно сделать следующие выводы:

- модель MF-VAR по данным смешанной частоты при наилучшем подборе высокочастотных переменных способна строить более точные прогнозы по сравнению с моделью VAR по агрегированным данным в режиме краткосрочного прогнозирования;
- на значения реального ВВП белорусской экономики в краткосрочной перспективе наибольшее влияние оказывают такие макроэкономические показатели, как объем промышленного производства, объем розничного товарооборота и индекс экономических настроений СИЭН;
- модель MF-VAR по данным смешанной частоты способна проводить качественный наукастинг в спокойных для экономики Беларуси периодах.

Таким образом, данное исследование вносит вклад в развитие эконометрических методов анализа временных рядов по данным смешанной частоты. По результатам исследования можно утверждать, что MF-VAR модель применима к белорусской экономике для краткосрочного прогнозирования и наукастинга с целью корректировки оценки ВВП.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Валовой внутренний продукт. [Электронный ресурс] – https://ru.wikipedia.org/wiki/%D0%92%D0%B0%D0%BB%D0%BE%D0%B2%D0%BE%D0%B9_%D0%B2%D0%BD%D1%83%D1%82%D1%80%D0%B5%D0%BD%D0%BD%D0%B8%D0%B9_%D0%BF%D1%80%D0%BE%D0%B4%D1%83%D0%BA%D1%82 – Дата доступа 27.05.2025.

2 Введение в анализ временных рядов : учебное пособие для вузов / Н. В. Артамонов, Е. А. Ивин, А. Н. Курбацкий, Д. Фантаццини ; Московский государственный университет имени М. В. Ломоносова, Московская школа экономики, Кафедра эконометрики и математических методов экономики. – Вологда : ВолНЦ РАН, 2021.

3 ВВП: что это, как определить и на что влияет. [Электронный ресурс] – <https://www.rbc.ru/quote/news/article/6273353d9a7947534ca8d991> – Дата доступа 27.05.2025.

4 Макеева, Н.М. Наукастинг элементов использования ВВП России / Н.М. Макеева, И.П. Станкевич // Статья 2022/10, Экономический журнал ВШЭ.

5 Камаев В. Д. Экономическая теория: учебник / под ред. В. Д. Камаева. — 10-е изд., перераб. и доп. — М.: ИНФРА-М, 2019. — 736 с.

6 Кругман П. Макроэкономика: учебник / П. Кругман, Р. Уэлс — М.: Издательство «Вильямс», 2018. — 624 с.

7 Манкью Н. Г. Принципы экономики: учебник / пер. с англ. — 9-е изд. — СПб.: Питер, 2020. — 832 с.

8 Макроэкономические модели. [Электронный ресурс] – <https://v1.savant.pro/community/3/~16/item/0ye69nzS> – Дата доступа 27.05.2025.

9 Макроэконометрические модели и их применение для оценки эффектов экономических политик. [Электронный ресурс] – <https://na-journal.ru/6-2024-ekonomika-menedzhment/12869-makroekonometricheskie-modeli-i-ih-primenenie-dlya-ocenki-eff> – Дата доступа 27.05.2025.

10 Малюгин, В. Краткосрочное прогнозирование и наукастинг темпов роста инфляции на основе моделей по смешанным данным / В.И. Малюгин // Банкаўскі веснік. – 2024. – С. 1–13.

11 Малюгин В., Крук Д., Милевский П. Индекс экономических настроений белорусской экономики: методические, модельные и программ-

ные средства. Журнал “Банковский вестник”. Исследования банка №16, 2019.

12 Понятие стационарности временного ряда. Процессы «единичного корня». [Электронный ресурс] – <https://bsu.by/upload/page/544153.pdf> – Дата доступа: 02.11.2024

13 Процессы «единичного корня». Тесты «единичного корня»: ADF, PP, KPSS. [Электронный ресурс] – <https://bsu.by/upload/page/546923.pdf> – Дата доступа: 02.11.2024

14 Станкевич И. П. Применение MIDAS-моделей с марковским переключением для наукастинга ВВП и его компонентов // Прикладная эко

15 Станкевич И.П. Сравнение методов наукастинга макроэкономических индикаторов на примере российского ВВП // Прикладная эконометрика 2020. С. 113–127.

16 Статистика инвестиций. [Электронный ресурс] – https://www.belstat.gov.by/upload-belstat/upload-belstat-word/Metod_pologenija/VVP_27_02_2018.doc – Дата доступа 05.04.2025.

17 Федеральная служба государственной статистики (Росстат). Методология расчёта валового внутреннего продукта (ВВП). – <https://rosstat.gov.ru/metodology/vvp> – Дата доступа 27.05.2025.

18 Харин, Ю. С. Теория вероятностей, математическая и прикладная статистика / Ю. С. Харин, Н. М. Зуев, Е. Е. Жук // Минск : БГУ, 2011.

19 Что такое номинальный ВВП. [Электронный ресурс] – <https://brobank.ru/nominalnyj-vvp/> – Дата доступа 27.05.2025.

20 Что такое индекс PMI. Объясняем простыми словами. [Электронный ресурс] – <https://secretmag.ru/enciklopediya/chto-takoe-indeks-pmi-obyasnyаем-prostymi-slovami.htm> – Дата доступа 28.05.2025.

21 Akaike information criterion. [Electronic resource] – https://en.wikipedia.org/wiki/Akaike_information_criterion – Access date 31.03.2025.

22 Almon, S. The distributed lag between capital appropriations and expenditures / S. Almon // Econometrica. – 1965. – № 33. – P. 178–196.

23 Andreou, E. Regression models with mixed sampling frequencies / E. Andreou, E. Ghysels, A. Kourtellos // Journal of Econometrics. – 2010. – № 58. – P. 246–261.

24 Augmented Dickey–Fuller test. [Electronic resource] – https://en.wikipedia.org/wiki/Augmented_Dickey_Fuller_test – Access date 31.03.2025.

- 25 Bayesian information criterion. [Electronic resource] – https://en.wikipedia.org/wiki/Bayesian_information_criterion – Access date 31.03.2025.
- 26 Dickey-Fuller test. [Electronic resource] – https://en.wikipedia.org/wiki/Dickey-Fuller_test – Access date 02.11.2024.
- 27 Ghysels, E. Midas regressions: Further results and new directions / E. Ghysels, A. Sinko, R. Valkanov // *Econometric Reviews*. – 2006. – № 26. – P. 53–90.
- 28 Ghysels, E. The MIDAS touch: Mixed data sampling regression models / E. Ghysels, P. Santa-Clara, R. Valkanov // Working paper, UNC and UCLA. – 2002. – 33 p.
- 29 Foroni, C. A comparison of mixed frequency approaches for nowcasting Euro area macroeconomic aggregates / C. Foroni, M. Marcellino // *International Journal of Forecasting*. – 2014a. – Vol. 30(3). – pp. 554–568. doi: 10.1016/j.ijforecast.2013.01.010
- 30 Foroni, C. A survey of econometric methods for mixed frequency data / C. Foroni, M. Marcellino // Working Paper 2013/06, Norges Bank. – 2013. – 45 p.
- 31 Foroni, C. U-MIDAS: MIDAS regressions with unrestricted lag polynomials / C. Foroni, Massimiliano M., Schumacher C. // Discussion Paper Series 1: Economic Studies 2011,35, Deutsche Bundesbank, 2011.
- 32 Hannan–Quinn information criterion. [Electronic resource] – https://en.wikipedia.org/wiki/Hannan-Quinn_information_criterion – Access date 31.03.2025.
- 33 IHS Markit. An introduction to the PMI surveys. IHS Markit, 2017.
- 34 Lutkepohl, H. New Introduction to Multiple Time Series Analysis // European University Institute, 1991.
- 35 Maccarrone G. GDP Forecasting: Machine Learning, Linear or Autoregression? / G. Maccarrone, G. Morelli, S. Spadaccini // *Front. Artif. Intell.* 4:757864, 2021.
- 36 Maravall A. An Application of TRAMO-SEATS: Model Selection and Out-of-sample Performance. / A. Maravall, F. J. Sanchez // Banco de Espana — Servicio de Estudios Documento de Trabajo n. 0014.
- 37 Parigi G., Golinelli R. The use of monthly indicators to forecast quarterly GDP in the short run: an application to the G7 countries // *Journal of Forecasting*. – 2007. – Vol. 26(2). – pp. 77–94. doi: 10.1002/for.1007

- 38 Parigi G., Schlitz G. Quarterly forecasts of the italian business cycle by means of monthly economic indicators. // Journal of Forecasting. – 1995. – Vol. 14(2). – pp. 117–141. doi: 10.1002/for.3980140205.
- 39 Perron, P. The Great Crash, the oil price shock, and the unit root hypothesis // Econometrica, 57, 1989.
- 40 Polasek W. Chow-Lin methods in spatial mixed models / W. Polasek, R. Sellner, Llano C. // The Rimini Centre for Economic Analysis, WP 10-47, 2010.
- 41 Runstler G., Sedillot F. Short-term estimates of euro area real GDP by means of monthly data // ECB Working Papers. – 2003. – N 276.
- 42 Trehan B. Forecasting growth in current quarter real GNP // Federal Reserve Bank of San Francisco Economic Review. – 1989. – pp. 39–42.
- 43 Unit root. [Electronic resource] – https://en.wikipedia.org/wiki/Unit_root – Access date 31.03.2025.