

Федеральное государственное автономное образовательное учреждение  
высшего профессионального образования «Национальный  
исследовательский университет «Высшая школа экономики»

Факультет экономических наук

Образовательная программа «Экономика»

Бакалаврская выпускная работа

**Сравнение прогнозной силы Байесовской векторной  
авторегрессии с одномерными моделями временных  
рядов**

Выполнила: Кузина Анна, БЭК139

Научный руководитель: Демешев Борис Борисович

Москва

2017 г.

# Содержание

|   |           |
|---|-----------|
| <b>Введение</b>   | <b>3</b>  |
| <b>1 Обзор Литературы</b>   | <b>6</b>  |
| 1.1 Одномерные модели временных рядов . . . . .                           | 6         |
| 1.2 Векторная авторегрессия . . . . .                                     | 8         |
| 1.3 Модели с переменными параметрами . . . . .                            | 9         |
| <b>2 Теоретическая часть</b>  | <b>11</b> |
| 2.1 Экспоненциальное сглаживание . . . . .                                | 11        |
| 2.2 ARIMA . . . . .   | 16        |
| 2.3 Тета-метод . . . . .  | 20        |
| 2.4 TVP-BVAR . . . . .  | 22        |
| 2.4.1 Модель . . . . .  | 22        |
| 2.4.2 Априорные распределения . . . . .                                   | 26        |
| 2.4.3 Алгоритм MCMC . . . . .   | 27        |
| 2.4.4 Прогнозирование . . . . .   | 30        |
| <b>3 Практическая часть</b>   | <b>33</b> |
| 3.1 Данные . . . . .  | 33        |
| 3.2 Методология . . . . .   | 36        |
| 3.2.1 Построение прогноза одномерных моделей . . . . .                    | 36        |
| 3.2.2 Построение прогноза TVP-BVAR . . . . .                              | 38        |
| 3.2.3 Сравнение прогнозов . . . . .                                       | 42        |
| 3.3 Результаты . . . . .  | 43        |
| 3.3.1 Результаты прогнозирования макроэкономических дан-<br>ных . . . . . | 44        |
| 3.3.2 Результаты прогнозирования финансовых данных . . . . .              | 47        |
| <b>Заключение</b>   | <b>51</b> |

|  |           |
|--|-----------|
| <b>Приложения</b>                      | <b>52</b> |
| А. Макроэкономические данные . . . . . | 52        |
| В. Финансовые данные . . . . .         | 53        |
| С. Расчеты . . . . .                   | 55        |

# Введение

Прогнозирование временных рядов заключается в предсказании их будущих значений по настоящим и прошлым значениям при помощи каких-либо научных методов. Прогнозирование применяется в большом количестве ситуаций, например, для финансового и торгового планирования, при принятии управленческих решений на предприятиях или при выборе оптимальной макроэкономической политики Центральными банками. В некоторых случаях требуется прогнозирование на несколько лет вперед, например, при выборе крупного проекта для инвестирования, а иногда — только на несколько минут, как при высокочастотной торговле на фондовом рынке.

Выделяют два основных типа методов прогнозирования — это качественные и количественные модели. К первым относится, например, использование экспертных оценок или проведение интервью. Однако, в текущей работе исследуются исключительно второй тип — количественные методы, основанные на предположении, что закономерности, наблюдаемые в данных, могут быть экстраполированы в будущее.

Целью данной работы является сравнение качества прогнозов, которые получаются при использовании байесовской векторной авторегрессии с переменными параметрами, с прогнозами более распространенных одномерных методов, таких как ARIMA, экспоненциальное сглаживание, тета-метод и методов гибридных прогнозов для горизонтов различной длины и временных рядов из разных источников. Для достижения указанной цели были поставлены следующие задачи:

- Сбор и первичная обработка финансовых и макроэкономических данных для последующего анализа

— Приведение данных к единому формату

- Проверка ряда на стационарность
- Разбиение выборки на тестовую и тренировочную
- Оценка моделей на тренировочной выборке и построение прогнозов
  - Выбор оптимальной спецификации модели
  - Оценка всех неизвестных параметров
  - Построение прогнозов будущих значений рядов
  - Расчет ошибок прогноза на тестовой выборке
- Сравнение полученных результатов для разных моделей

Как уже было сказано выше, задача прогнозирования временных рядов имеет высокое значение для различных сфер экономики, а в силу большого числа доступных статистических методов не менее важной задачей становится выбор модели, которая лучше всего справится с прогнозированием. В силу ограниченности ресурсов, на практике выбор из огромного множества доступных моделей осуществляется на основе результатов, которые данная модель продемонстрировала в различных исследованиях. Именно поэтому необходимы работы, в которых новые методы сравниваются с более привычными или простыми моделями на различных данных.

Байесовская векторная авторегрессия достаточно часто применяется при прогнозировании макроэкономических показателей, так как позволяет включать в модель большое количество переменных и их лагов, а также учитывать априорные представления о распределении параметров модели. Использование моделей с переменными параметрами дает еще более широкие возможности для анализа совместной динамики переменных, хотя и значительно усложняет процедуру оценки модели. В данной работе исследуется именно такая модель и используется относительно новый алгоритм семплирования из статьи Primiceri, Del Negro (2015), призванный устранить ошибку в алгоритме, который использовался в более ранних работах.

Первая часть работы посвящена обзору литературы, относящейся к выбранным методам прогнозирования. Во второй части подробно описаны используемые модели, их возможные спецификации и способы оценки параметров. Третья часть описывает данные, которые были выбраны для проведения анализа, методологию построения прогнозов и их сравнения между собой, а также полученные результаты.

# 1. Обзор Литературы

Ранние подходы к изучению временных рядов базировались на предположении об их детерминированности. То есть исследователи считали, что информация о временном ряде на некотором временном горизонте позволяет точно судить о его поведении вне этого интервала. Переход к стохастическому подходу произошел только в XX веке. Одной из поворотных статей в данной области стала работа Yule (1927), где временной ряд рассматривался как реализация некоторого случайного процесса. Впоследствии, основываясь на этой простой идее, ученые разработали огромное количество методов моделирования временных рядов.

## 1.1. Одномерные модели временных рядов

Наиболее широко используемыми подходами к прогнозированию одномерных временных рядов являются экспоненциальное сглаживание и модели ARIMA (autoregressive integrated moving average). Рассмотрим подробнее основную литературу, посвященную данным методам.

Модели экспоненциального сглаживания берут начало в литературе 1950-х годов. В статьях Holt (1957), Brown (1959) и Winters (1960) впервые использовался прогноз, представляющий собой среднее взвешанное всех доступных наблюдений с экспоненциально убывающими весами. В статье Muth (1960) впервые обсуждаются статистические свойства временных рядов, которые описываются моделью простого экспоненциального сглаживания. Автор показал, что такая модель является оптимальной для прогнозирования случайного блуждания с шумом.

В дальнейшем, большое количество работ было посвящено возможным расширениям простого экспоненциального сглаживания, самые популярные из которых — метод Хольта с аддитивным трендом, метод Хольта-Уинтерса с аддитивным трендом и сезонностью (аддитивной или мульт-

типликативной). В статье Hyndman, Koehler, Snyder, and Grose (2002) подробно рассматриваются все возможные спецификации моделей экспоненциального сглаживания, их формулировки в виде моделей пространства состояний с мультипликативными и аддитивными ошибками, а также сравнивается качество прогнозов, полученных в результате применения разных спецификаций к различным временным рядам.

Еще один крайне популярный способ прогнозирования временных рядов, ARIMA, представляет собой комбинацию авторегрессии (AR) и модели скользящего среднего (MA), допускающую взятие разности исходного ряда. Внушительное число работ было посвящено выбору спецификаций модели, оценке коэффициентов и прогнозированию с использованием моделей из семейства ARIMA. В результате, все основные, доступные на тот момент, научные достижения в этой области были объединены в работе Box & Jenkins (1970). В своей книге авторы подробно описывают разработанный ими подход к анализу временных рядов, известный до сих пор как методология Бокса-Дженкинса, который состоит из выбора подходящей модели, оценки параметров и проверки соблюдения всех необходимых предпосылок.

По мере появления все большего числа моделей для прогнозирования временных рядов, исследователи стали задумываться об эффективности комбинации прогнозов различных моделей. Одной из первых работ на эту тему считается статья Bates and Granger (1969), где авторы представили несколько способов комбинации точечных прогнозов и показали, что такая техника способна улучшить качество прогноза по сравнению с результатами каждого метода по отдельности. На данный момент существует огромное количество способов комбинации прогнозов. Например, в статье Granger and Ramanathan (1984) было предложено оценивать веса прогнозов, используя метод наименьших квадратов на прошлых данных, а Deutsch, Granger, and Terasvirta (1994) предлагали использовать веса, изме-



няющиеся во времени.

## 1.2. Векторная авторегрессия

Для построения многомерных моделей одним из наиболее популярных инструментов является векторная авторегрессия (VAR). Представленная впервые в работе Sims (1980), данная модель отличается относительной простотой и не требует введения никаких ограничений на совместную динамику временных рядов. У VAR моделей существует большое количество спецификаций, благодаря чему их можно использовать для прогнозирования различных видов данных. Так, в статье Funke (1990) сравниваются результаты прогноза одномерной модели ARIMA и пяти разных спецификаций VAR модели (в частности структурной и байесовской VAR с различными априорными распределениями). Основной проблемой VAR моделей часто называют излишнюю параметризацию, которая может возникнуть при увеличении числа размерности модели или с ростом числа включаемых лагов и привести к невозможности получения оценок коэффициентов.

Для решения описанной выше проблемы Litterman (1986) предложил ввести ограничение в виде априорного распределения параметров. Модели, в которых используется байесовский подход к оценке коэффициентов получили название BVAR (Bayesian Vector Autoregression). Главный недостаток байесовских моделей — необходимость использования сопряженных распределений, перестал быть актуальным в современной науке благодаря высокой скорости, с которой современные компьютеры проводят симуляции и, как следствие, возможности применения методов MCMC (Markov Chain Monte Carlo). Если же говорить о плюсах данного подхода, то он, в первую очередь, позволяет включить в модель большое число переменных и лагов, не опасаясь излишней параметризации. Более того, выбирая из широкого круга априорных распределений, исследователи имеют

возможность снизить неопределенность относительно значений параметров, опираясь на опыт коллег или общепринятые представления о динамике тех или иных переменных.

Существует огромное количество работ, в которых BVAR используют для прогнозирования всевозможных временных рядов. Так, можно привести примеры прогнозирования макроэкономических данных (Artis and Zhang (1990); Beauchemin, Kenneth and Zaman (2011); Berg, Oliver and Henzel (2013)), рыночных долей компаний (Ramos (2003)), финансовых рядов (Carriero, Kapetanios and Marcellino (2012))

### **1.3. Модели с переменными параметрами**

Модели, в которых параметры изменяются со временем, достаточно широко используются в макроэкономике и финансах. Такая тенденция вполне объяснима, так как закономерностям в данных отраслях свойственно изменяться по мере развития технологического прогресса, изменения конъюнктуры и перемен в обществе. Одной из первых работ, в которой байесовская векторная авторегрессия применялась для прогнозирования временных рядов была Doan, Litterman and Sims (1984). Авторы применили фильтр Калмана для обновления вектора параметров, представляющих собой процесс  $AR(1)$  и оценили значения гиперпараметров, при которых их модель давала наилучший прогноз среди других VAR моделей.

По мере роста популярности байесовских методов стал расти и интерес к векторным авторегрессиям с переменными коэффициентами (TVP-VAR). Наиболее знаменитые работы в этой области — Cogley and Sargent (2002, 2005), Primiceri (2005), Koop and Korobilis (2009). В первых трех статьях прогнозированию не уделялось особого значения, однако, в них были предложены основные принципы применения байесовского подхода к TVP-VAR моделям.

В статье Primiceri (2005) была подробно рассмотрена байесовская векторная авторегрессия с переменными параметрами (далее — TVP-BVAR). У данной модели существует две основные особенности: во-первых, предполагается, что все коэффициенты модели могут изменяться со временем, а во-вторых, ковариационная матрица остатков также зависит от времени. Такие предположения позволяют строить модель для длительных временных промежутков, подстраиваясь под изменения конъюнктуры и меняющиеся зависимости между рассматриваемыми переменными. Как уже было сказано выше, автор статьи описывает применение байесовского подхода к подобной модели, а также предлагает алгоритм, позволяющий получить выборку из апостериорного распределения параметров. Однако, через несколько лет после публикации данной статьи в алгоритме была обнаружена ошибка, которую авторы исправили относительно недавно, опубликовав статью Primiceri, Del Negro (2015). Вскоре после этого была опубликована статья Kruger (2015), где был представлен пакет `bvarsv` для языка программирования R, позволяющий оценивать модель TVP-BVAR с использованием актуального алгоритма.

## 2. Теоретическая часть

Рассмотрим одномерные модели, которые были выбраны в качестве конкурентов для основной модели работы — байесовской векторной авто-регрессии с переменными коэффициентами и стохастической волатильностью. Этими одномерными моделями являются экспоненциальное сглаживание, ARIMA, тета-метод и их комбинации (гибрид из нескольких одномерных моделей). Подробнее процедура формирования гибридных моделей будет описана в разделе [3.2](#), посвященному методологии работы.

### 2.1. Экспоненциальное сглаживание

Семейство моделей экспоненциального сглаживания достаточно часто используются для прогнозирования временных рядов. Можно предположить, что они стали популярными, отчасти, благодаря своей простоте и гибкости. В основе всех моделей данного семейства лежит простое экспоненциальное сглаживание — метод прогнозирования временных рядов, базирующийся на предположении, что каждое последующее наблюдение временного ряда может быть представлено как взвешенная сумма всех прошлых наблюдений. Причем, чем дальше наблюдение от последнего доступного, тем меньше его вес. Таким образом, последовательность весов представляет собой убывающую геометрическую прогрессию. Если записать описанный выше метод в виде математического выражения, то мы получим следующее уравнение для определения прогноза временного ряда:

$$\begin{aligned}\hat{y}_{T+1|T} &= \alpha y_T + \alpha(1 - \alpha)y_{T-1} + \alpha(1 - \alpha)^2 y_{T-2} + \dots = \\ &= \alpha y_T + (1 - \alpha)\hat{y}_{T|T-1} \quad (1)\end{aligned}$$

Так как уравнение (1) позволяет получить прогноз только на 1 период вперед, для любого горизонта  $h > 1$  прогноз будет равен одному и тому же значению, прогнозу для  $h = 1$ :

$$\hat{y}_{T+h|T} = \hat{y}_{T+1|T}$$

Если записать уравнение (1) для каждого доступного наблюдения, а потом подставить выражение для каждого сглаженного значения в предшествующее ему уравнение мы получим следующее выражение для прогноза на один период:

$$\hat{y}_{T+1|T} = \sum_{i=0}^{T-1} \alpha(1-\alpha)^i y_{T-i} + (1-\alpha)^T \hat{y}_{1|0}$$

В результате, для того, чтобы получить экспоненциально сглаженный ряд, нам необходимо задать два параметра: начальное значение сглаженного ряда  $\hat{y}_{1|0}$  (прогноз первого доступного наблюдения) и параметр  $\alpha$ .

Если есть необходимость получить не только точечные прогнозы, но и их доверительные интервалы, нам необходимо определиться со статистической моделью, которая (как мы предполагаем) описывает наш временной ряд  $y_t$ . В случае с экспоненциальным сглаживанием можно построить модель, состоящую из двух основных компонентов: уравнения измерения (measurement equation), описывающего наблюдаемые данные, и уравнений перехода (transition equations), которые описывают ненаблюдаемые компоненты данных (например, тренд или сезонность). Модели, состоящие из таких элементов называют моделями пространства состояний (state space models).

Все что нужно, чтобы перейти от детерминированных уравнений, описанных выше, к стохастической модели пространства состояний, это задать распределение ошибок  $\varepsilon_t$ . Предположив, что ошибки независимы и

нормальны, то есть  $\varepsilon_t = y_t - \hat{y}_{t|t-1} \sim i.i.d.N(0, \sigma^2)$ , а также обозначив сглаженное значение (уровень ряда)  $l_{t-1} = \hat{y}_{t|t-1}$ , получим следующую модель:

$$\begin{aligned} y_t &= l_{t-1} + \varepsilon_t \\ l_t &= l_{t-1} + \alpha \varepsilon_t \end{aligned} \tag{2}$$

Если же вместо обычной ошибки прогноза принять  $\varepsilon_t$  равным относительной ошибке, то мы получим модель с мультипликативными ошибками, которая имеет немного другой формат записи.

$$\begin{aligned} y_t &= l_{t-1}(1 + \varepsilon_t) \\ l_t &= l_{t-1}(1 + \alpha \varepsilon_t), \\ \text{где } \varepsilon_t &= \frac{y_t - \hat{y}_{t|t-1}}{\hat{y}_{t|t-1}} \sim i.i.d.N(0, \sigma^2) \end{aligned}$$

Заметим, однако, что значение точечных прогнозов не зависит от вида ошибок, то есть при использовании модели с аддитивными и мультипликативными ошибками будут получаться совершенно одинаковые точечные прогнозы. Отличие будет заметно только за счет разных доверительных интервалов.

Стоит заметить, что простое экспоненциальное сглаживание может показывать очень плохие результаты, особенно когда оно имеет дело с данными, обладающими трендом или сезонностью. Именно по этой причине существуют другие, более продвинутое версии данной модели. Начнем с метода Хольта, который используется для прогнозирования данных с линейным трендом. В простейшем случае предполагается, что тренд входит в модель аддитивно, то есть появляется новый компонент в уравнении измерений, а также добавляется еще одно уравнение переходов, описывающее

изменение тренда во времени:

$$\begin{aligned}y_t &= l_{t-1} + b_{t-1} + \varepsilon_t \\l_t &= l_{t-1} + b_{t-1} + \alpha \varepsilon_t \\b_t &= b_{t-1} + \beta \varepsilon_t\end{aligned}\tag{3}$$

В данном случае прогноз уже не будет горизонтальным, так как к уравнению уровня добавится линейный тренд:

$$y_{T+h|T} = l_T + hb_t$$

Таким образом, для получения прогноза нам нужно определить оптимальное значение параметров  $\alpha, \beta$ , а также начальные значения тренда и уровня —  $l_0, b_0$ . Помимо такой спецификации, выделяют также модель с экспоненциальным трендом. В этом случае компонента  $b_t$  входит в уравнение измерений мультипликативно. Более того, существуют модели с децифированным трендом (damped trend), в которых по мере увеличения горизонта прогноза снижается и вклад трендовой составляющей, что позволяет получать более правдоподобные долгосрочные прогнозы.

Помимо тренда, временные ряды часто обладают свойством сезонности, которую крайне важно учитывать при прогнозировании. Модель Хольта-Уинтерса позволяет получить экспоненциально сглаженный ряд с учетом наличия аддитивной или мультипликативной сезонности в данных. Добавление сезонности в модель, так же как и в случае с трендом, всегда означает появление нового компонента в уравнении измерений (входящего аддитивно или мультипликативно), а также нового уравнения переходов. Для примера рассмотрим, как изменится модель с линейным трендом (3),

если в нее добавить мультипликативную сезонность.

$$y_t = (l_{t-1} + b_{t-1})s_{t-m} + \varepsilon_t$$

$$l_t = l_{t-1} + b_{t-1} + \alpha \varepsilon_t / s_{t-m}$$

$$b_t = b_{t-1} + \beta \varepsilon_t / s_{t-m}$$

$$s_t = s_{t-m} + \gamma \varepsilon_t / (l_{t-1} + b_{t-1})$$

В данной работе, при прогнозировании методом экспоненциального сглаживания, используется функция `ets()` пакета `forecast`, написанного для языка программирования R. Для получения оценок параметров модели и начальных значений тренда и сезонности (если это необходимо) в пакете используется метод максимального правдоподобия, максимизирующий вероятность получения наблюдаемой выборки при заданных ограничениях. Ограничения задаются таким образом, чтобы сглаживающие параметры уровня, тренда и сезонности лежали в пределах  $(0, 1)$ .

Заметим, что комбинируя различные виды тренда и сезонности можно получить 15 различных моделей, в каждой из которых могут быть мультипликативные или аддитивные ошибки. Классификация всех возможных комбинаций может быть представлена в виде таблицы.

Таблица 1: Классификация моделей экспоненциального сглаживания

| Тренд                            | Сезонность  |            |                   |
|----------------------------------|-------------|------------|-------------------|
|                                  | Отсутствует | Аддитивная | Мультипликативная |
| Отсутствует                      | $NN$        | $NA$       | $NM$              |
| Аддитивный                       | $AN$        | $AA$       | $AM$              |
| Аддитивный демпфированный        | $A_dN$      | $A_dA$     | $A_dM$            |
| Мультипликативный                | $MN$        | $MA$       | $MM$              |
| Мультипликативный демпфированный | $M_dN$      | $M_dA$     | $M_dM$            |

Источник: Hyndman, Koehler, Snyder, and Grose (2002), стр. 440



Для выбора оптимальной спецификации используется информационный критерий Акаике, скорректированный на размер выборки, формула для расчета которого представлена в уравнении (4). Оптимальной считается та модель, значение информационного критерия для которой минимально.

$$AIC_c = -2 \log(L) + 2k + \frac{2k(k+1)}{T-k-1}, \quad (4)$$

где  $L$  — значение функции правдоподобия, а  $k$  — общее число коэффициентов в модели (включая начальные значения уровня, тренда и сезонности, если они есть).

## 2.2. ARIMA

ARIMA представляет собой комбинацию авторегрессии и модели скользящего среднего. Полная спецификация модели  $ARIMA(p, d, q)$ , где  $p$  — порядок авторегрессионной части,  $q$  — порядок части скользящего среднего и  $d$  — порядок разности ряда, выглядит следующим образом:

$$\Delta : {}^d y_t = \alpha + \beta_1 \Delta^d y_{t-1} + \dots + \beta_p \Delta^d y_{t-p} + \varepsilon_t + \gamma_1 \varepsilon_{t-1} + \dots + \gamma_q \varepsilon_{t-q} \quad (5)$$

Можно записать модель с помощью оператора лага  $L$ :

$$(1 - L)^d (1 - \beta_1 L - \dots - \beta_p L^p) y_t = \alpha + (1 + \gamma_1 L + \dots + \gamma_q L^q) \varepsilon_t$$

Существует также разновидность моделей ARIMA, использующаяся для моделирования сезонных данных. Сезонная модель получается в результате добавления сезонных разностей в исходную модель (5). Такая модель обычно обозначается как  $ARIMA(p, d, q)(P, D, Q)_m$ , где  $m$  — частотность ряда (например,  $m = 4$  для квартальных данных). Она также может быть записана при помощи оператора лага. В результате, наиболее

обобщенная версия модели ARIMA имеет следующий вид:

$$\begin{aligned} (1 - L)^d(1 - L^m)^D(1 - \beta_1 L - \dots - \beta_p L^p)(1 - B_1 L^m - \dots - B_P L^{Pm})y_t = \\ = \alpha + (1 + \gamma_1 L + \dots + \gamma_q L^q)(1 + \Gamma_1 L^m + \dots + \Gamma_Q L^{Qm})\varepsilon_t \end{aligned}$$

Для определения параметра  $d$  можно провести тест на единичный корень. В функции `auto.arima()` используется KPSS тест (Kwiatkowski, Phillips, Schmidt, Shin), с помощью которого определяется минимальное число разностей. Авторы пакета в своей статье Hyndman and Khandakar (2008) утверждают, что для автоматической оценки модели гораздо эффективнее использовать тесты на единичный корень, в которых нулевая гипотеза предполагает отсутствие единичного корня, так как тесты с нулевой гипотезой о существовании единичного корня ведут к смещенным результатам. KPSS тест предполагает, что рассматриваемый временной ряд имеет следующий вид:

$$\begin{aligned} y_t &= c_t + \beta t + u_t \\ c_t &= c_{t-1} + \varepsilon_t, \text{ где } \varepsilon_t \sim WN(0, \sigma^2) \end{aligned}$$

Такой ряд является стационарным, если  $c_t$  постоянен во времени, то есть дисперсия  $\sigma^2 = 0$  нулевая и альтернативная гипотезы формулируются следующим образом:

$$\begin{aligned} H_0 : \sigma^2 &= 0 \quad (\text{временной ряд стационарен}) \\ H_a : \sigma^2 &\neq 0 \quad (\text{временной ряд не является стационарным}) \end{aligned}$$

Тестовая статистика для проверки выдвинутой гипотезы считается следующим образом:

$$KPSS = \frac{\sum_{i=1}^T S_t^2}{T^2 \lambda^2},$$

где  $S_t = \sum_{j=1}^t \hat{u}_j$  — сумма остатков в регрессии, а  $\lambda^2$  — состоятельная оценка дисперсии  $u_t$  в форме Ньюи-Уэста. Сравнив полученное значение тестовой статистики с асимптотическим распределением, можно сделать вывод о правдоподобности нулевой гипотезы. Если гипотеза о стационарности отвергается, то берется первая разность исходного ряда и тест повторяется еще раз. Оптимальное значение параметра  $d$  определяется как минимальное число разностей ряда, необходимое для того, чтобы нулевая гипотеза о его стационарности не отвергалась.

Для определения порядка сезонной разности в функции `auto.arima()` используется OSCB тест. Данный тест, представленный в статье Osborn, Chui, Smith, Birchenhall (1988), заключается в проверке нулевой гипотезы о наличии сезонного единичного корня. Он был выбран авторами пакета, так как позволил получить модели с лучшими прогнозами, по сравнению с остальными тестами для определения параметра  $D$ . Допустимые значения сезонной разности ограничены множеством  $\{0, 1\}$ , в результате чего тест на сезонный единичный корень проводится только один раз.

Определив параметры  $d$  и  $D$  описанными выше способами, функция `auto.arima()` выбирает оптимальный порядок  $p, q, P, Q$ , минимизируя скорректированный на размер выборки информационный критерий Акаике, который аналогичен данному критерию из моделей экспоненциального сглаживания (см. уравнение (4)). Число коэффициентов в модели (включая дисперсию остатков) можно определить следующим образом:

$$k = p + q + c + 1,$$

где  $c$  — дамми, отвечающая за наличие свободного коэффициента в модели.

Стоит, однако, заметить, что даже если выбирать из ограниченного числа параметров, количество моделей, для которых необходимо посчи-

тать значение критерия Акаике, очень велико. Чтобы процесс автоматического выбора модели не требовал огромных вычислительных мощностей, авторы пакета предложили использование ступенчатого подхода, состоящего из следующих шагов:

**Шаг 1** Оцениваются 4 модели:

- $\text{ARIMA}(2, d, 2)(1, D, 1)_m$ ;
- $\text{ARIMA}(0, d, 0)(0, D, 0)_m$  ;
- $\text{ARIMA}(1, d, 0)(1, D, 0)_m$ ;
- $\text{ARIMA}(0, d, 1)(0, D, 1)_m$ .

Отметим, что если  $m = 1$ , то рассматриваются модели без сезонной составляющей. Свободный коэффициент  $\alpha$  включается в модель только в том случае, если  $d + D \leq 1$ . Из четырех указанных моделей выбирают «текущую» модель, которая характеризуется наименьшим значением  $AIC_c$ .

**Шаг 2** Рассматриваются 13 модификаций текущей модели, :

- один из параметров  $p, q, P, Q$  отклоняется на 1 от текущего значения;
- параметры  $p$  и  $q$  одновременно отклоняются на 1 от текущего значения;
- параметры  $P$  и  $Q$  одновременно отклоняются на 1 от текущего значения;
- в модель включается свободный коэффициент, если его не было в текущей модели, или оценивается модель без свободного коэффициента, если в текущей модели он был.

Если в результате данных модификаций была найдена модель с меньшим значением  $AIC_c$ , то она становится текущей и шаг 2 повторяется еще раз. Если у всех модифицированных моделей значение выше текущего, то процесс выбора завершается.

Наконец, при заданных  $p, d, q$ , методом максимального правдоподобия определяются параметры  $\alpha, \beta_i, \gamma_i$ .

В результате, используя все полученные значения параметров, рассчитывается прогноз временного ряда из уравнения (5). Для этого нужно только выразить  $y_{T+h}$ , подставив  $T + h$  вместо  $t$ , а также заменив все будущие значения их прогнозами, будущие ошибки — нулями, а прошлые ошибки — соответствующими остатками. Очевидно, что можно получать прогнозы только последовательно, например, для прогнозирования наблюдения  $T + 3$ , где  $T$  — текущий момент времени, нам необходимо иметь прогноз для двух предыдущих периодов, чтобы подставить их на место ненаблюдаемых  $y_{T+2}$  и  $y_{T+1}$ .

### 2.3. Тета-метод

Представленный в работе Assimakopoulos, Nikolopoulos (2000), данный метод основан на прогнозировании с использованием декомпозиции. Однако, в отличие от классического подхода к декомпозиции, включающего в себя выделение сезонной компоненты, тренда и цикличности, в тета-методе производится выделение краткосрочной и долгосрочной компоненты. Прогнозом временного ряда в таких моделях является комбинация прогнозов каждой из компонент.

Упомянутые выше компоненты получаются в результате следующего преобразования исходного ряда  $\{y_t\}$ :

$$\Delta^2 \hat{y}_{t,\theta} = \theta \Delta^2 y_t, \quad \text{где}$$

$$\Delta^2 y_t = y_t - 2y_{t-1} + y_{t-2}$$

Подставляя разные значения параметра  $\theta$ , можно получать различные, так называемые, тета-линии. Так, если установить этот параметр равным 0, мы получим прямую, которую авторы статьи рассматривают как долгосрочную компоненту исходного ряда. При  $\theta > 1$  большую роль играют краткосрочные колебания данных, поэтому в оригинальной работе тета-линия с параметром 2 используется в качестве краткосрочной компоненты ряда. Получить прогноз первой, долгосрочной, компоненты, не доставляет труда, так как мы можем просто экстраполировать полученный линейный тренд. В случае со второй компонентой, авторы предлагают использовать для прогнозирования простое экспоненциальное сглаживание. Далее полученные прогнозы комбинируют, чтобы получить прогноз исходного ряда. В результате, прогноз на  $h$  периодов вперед принимает следующий вид:

$$y_{T+h|T} = \frac{1}{2}(\hat{y}_{T+h,0} + \hat{y}_{T+h,2})$$

В статье Hyndman and Billah (2008) была предложена стохастическая модель пространства состояний для данного метода прогнозирования.

$$y_t = l_{t-1} + b + \varepsilon_t$$

$$l_t = l_{t-1} + b + \alpha \varepsilon_t,$$

где  $\varepsilon_t \sim i.i.dN(0, \sigma^2)$ , а начальное состояние задается как  $y_1 = l_1$ .

Пакет `forecast`, как и в случае с описанными выше одномерными моделями, предоставляет возможность автоматической оценки данной модели. Параметр  $\alpha$ , необходимый для оценки тета-линии с параметром  $\Theta = 2$ , определяется при помощи информационного критерия Акаике (см. уравнение (4)).

## 2.4. TVP-BVAR

### 2.4.1. Модель

Наконец, рассмотрим байесовскую векторную авторегрессию с переменными коэффициентами и стохастической волатильностью, описанную в статье Primiceri (2005). Как видно из названия, она представляет из себя векторную авторегрессию, в которой коэффициенты и ковариационная матрица остатков изменяются со временем и оценки неизвестных параметров находятся посредством байесовского подхода. Построение такого рода моделей позволяет учитывать нелинейности зависимостей между разными переменными или их временные изменения.

Обозначим вектор из  $n$  эндогенных переменных как  $y_t$  и рассмотрим основное уравнение модели:

$$y_t = c_t + B_{1,t}y_{t-1} + \dots + B_{k,t}y_{t-k} + \varepsilon_t, \quad t = 1, \dots, T, \quad (6)$$

где  $c_t$  —  $n \times 1$  вектор свободных коэффициентов, меняющийся во времени,  $B_{i,t}, i = 1, \dots, p$  —  $n \times n$  матрицы коэффициентов, также зависящих от момента времени,  $u_t$  — гетероскедастичные остатки с ковариационной матрицей  $\Omega_t$ . Авторы приводят ковариационную матрицу остатков к диагональному виду, используя следующее преобразование:

$$A_t \Omega_t A_t' = \Sigma_t \Sigma_t' \quad (7)$$

Здесь матрица  $A_t$  является нижнетреугольной и содержит в себе параметры, отвечающие за связь между ортогонализированными остатками модели, а  $\Sigma_t$  является диагональной и состоит из стандартных ошибок ортогонализированных остатков:

$$A_t = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \alpha_{21,t} & 1 & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \alpha_{n1,t} & \dots & \alpha_{nn-1,t} & 1 \end{bmatrix} \quad \Sigma_t = \begin{bmatrix} \sigma_{1,t} & 0 & \dots & 0 \\ 0 & \sigma_{2,t} & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_{n,t} \end{bmatrix}$$

При таком разложении шоки модели можно выразить как:

$$u_t = A^{-1} \Sigma_t \varepsilon_t, \text{ где } \text{Var}(\varepsilon_t) = I_n$$

Не сложно показать, что дисперсия шоков осталась прежней:

$$\text{Var}(u_t) = A^{-1} \Sigma_t \text{Var}(\varepsilon_t) (\Sigma_t^{-1})' = A^{-1} \Sigma_t \Sigma_t' (A_t^{-1})' = \Omega_t$$

В итоге, мы можем рассмотреть модель пространства состояний, содержащую уравнение измерений (6) и три уравнения переходов: для параметров  $B_t$  и значений матрицы  $A_t$  (не равных 0 или 1), а также для диагональных элементов матрицы  $\Omega_t$ . Пусть  $\alpha_t$  — вектор всех переменных элементов матрицы  $A_t$ , а  $\sigma_t$  — вектор диагональных элементов  $\Omega_t$ . Тогда полная модель имеет следующий вид:

$$y_t = X_t' B_t + A_t^{-1} \Omega_t \varepsilon_t \quad (8)$$

$$B_t = B_{t-1} + \nu_t, \quad \text{где } \nu_t \sim N(0, Q) \quad (9)$$

$$\alpha_t = \alpha_{t-1} + \zeta_t, \quad \text{где } \zeta_t \sim N(0, S) \quad (10)$$

$$\log \sigma_t = \log \sigma_{t-1} + \eta_t, \quad \text{где } \eta_t \sim N(0, W) \quad (11)$$



Отметим, что уравнение (8) отличается от (6) только тем, что оно записано в матричной форме, из чего следует что  $X'_t = I_n \otimes [1, y'_{t-1}, \dots, y'_{t-k}]$  — матрица размера  $n \times n(k+1)$ , а  $B_t$  — вектор размерности  $n(k+1) \times 1$ , состоящий из всех коэффициентов  $\{B_{i,t}\}_{i=1}^k$  и свободных коэффициентов  $c_t$  в момент времени  $t$ .

Автор модели предполагает, что элементы вектора  $B_t$  и переменные коэффициенты матрицы  $A_t$  изменяются по закону случайного блуждания, в то время как стандартные отклонения  $\sigma_t$ , входящие в модель мультипликативно, моделируются с помощью процесса геометрического случайного блуждания.

Наконец, предполагается, что случайные величины  $\{\varepsilon_t, \nu_t, \zeta_t, \eta_t\}$  независимы. Авторы подчеркивают, что данное ограничение не является критически важным для построения данной модели, однако, его использование значительно снижает количество коэффициентов и упрощает интерпретацию полученных результатов. Более того, предполагается, что матрица  $S$  имеет блочно-диагональную структуру, где каждый блок отвечает за параметры  $\alpha_{i,t}$  одного уравнения. Так, если рассмотреть модель для трех переменных, то вектор  $\alpha_t$  будет состоять из трех элементов:  $(\alpha_{21,t}, \alpha_{31,t}, \alpha_{32,t})$ , каждый из которых является процессом случайного блуждания (10). Тогда вектор шоков имеет ковариационную матрицу  $S$ , имеющую следующую структуру:

$$S = \text{Var} \begin{pmatrix} \zeta_{21,t} \\ \zeta_{31,t} \\ \zeta_{32,t} \end{pmatrix} = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}, \text{ где}$$

$$S_1 = \text{Var}(\zeta_{21,t})$$

$$S_2 = \begin{pmatrix} \text{Var}(\zeta_{31,t}) & \text{Cov}(\zeta_{31,t}, \zeta_{32,t}) \\ \text{Cov}(\zeta_{31,t}, \zeta_{32,t}) & \text{Var}(\zeta_{32,t}) \end{pmatrix}$$

Такое ограничение подразумевает, что коэффициенты одновременных соотношений переменных в модели (вектор  $\alpha_t$ ), изменяются независимо в каждом уравнении модели. Как и предположение о независимости шоков различных параметров между собой, данная предпосылка является некоторым упрощением, которое снижает количество вычислений, необходимых для оценки параметров. При этом авторы показывают, что данные упрощения не оказывают значительного влияния на эмпирические результаты, что свидетельствует об оправданности их использования.

Количество параметров, которые требуется оценить в модели (8) – (11), обычно достигает нескольких тысяч. Так, в модели с 3 переменными и двумя лагами содержится 21 коэффициент, каждый из которых изменяется со временем (что при 100 периодах уже означает наличие 2100 коэффициентов), а число гиперпараметров (свободные элементы матриц  $Q, S, W$ ) составляет 241<sup>1</sup>. Оценка такого количества параметров возможна только байесовским методом. Рассмотрим подробнее принцип его работы.

Байесовский подход заключается в том, что имея некоторую модель и априорные представления о ее параметрах, можно использовать формулу Байеса для условной вероятности, чтобы обновлять эти представления, используя доступные данные. Распределение, получаемое в результате, называется апостериорным. Именно его используют как источник информации о параметрах модели. Ниже представлена формула Байеса, которая показывает, как апостериорное распределение зависит от априорного и функции правдоподобия.

$$P(\Theta|Y_T) = \frac{P(Y_T|\Theta) P(\Theta)}{P(Y_T)} \propto \text{Prior} \cdot \text{Likelihood} \quad (12)$$

---

<sup>1</sup>Primiceri 2005, стр. 28

## 2.4.2. Априорные распределения

В данном разделе будут рассмотрены априорные распределения гиперпараметров модели TVP-BVAR и начальных значений процессов случайного блуждания (9) – (11). В оригинальной статье Primiceri предполагается, что все начальные значения независимы и имеют нормальное распределение, а матрицы гиперпараметров имеют обратное распределение Уишарта и также независимы друг от друга (в том числе и блоки матрицы  $S$ ). Рассмотрим подробнее все априорные распределения, необходимые для оценки параметров модели.

- $B_0 \sim N(B_{OLS}, k_B \times \widehat{\text{Var}}(B_{OLS}))$

Таким образом, начальные значения коэффициентов распределены нормально с математическим ожиданием, равным их МНК оценкам. Для получения этих значений авторы статьи, а впоследствии и автор пакета для R, используют первые 40 наблюдений тренировочной выборки. Значение параметра  $k_B$  принимается равным 4.

- $A_0 \sim N(A_{OLS}, k_A \times \widehat{\text{Var}}(A_{OLS}))$

Для получения математического ожидания начального значения процесса  $\alpha_t$  строятся линейные регрессии остатков модели из предыдущего пункта друг на друга. То есть оцениваются регрессии следующего вида:

$$\hat{u}_{i,t} = \alpha_{ij,t|OLS} \cdot \hat{u}_{j,t} + \hat{e}_t,$$

где  $\hat{u}_i$  — остаток регрессии  $i$ -ой переменной на свои лаги и лаги всех остальных переменных. Параметра  $k_A$ , как и в предыдущем пункте, полагается равным 4.

- $\log(\sigma_0) \sim N(\log(\sigma_{OLS}), I_n)$

В данном случае МНК оценка стандартного отклонения может быть

получена из остатков регрессий предыдущего пункта посредством разложения (7).

- $Q \sim IW(k_Q^2 \times p_Q \times \widehat{\text{Var}}(B_{OLS}), p_Q)$

В оригинальной статье используются следующие параметры:  $p_Q$  задается равным размеру обучающей выборке для построения МНК оценок, то есть 40, а  $k_Q$  выбрано равным 0.01.

- $W \sim IW(k_W^2 \times p_W \times I_n, p_W)$

Количество степеней свободы,  $p_W$ , выбирается в зависимости от количества переменных в модели и равно  $n + 1$ , а  $k_W = 0.01$ , как и в предыдущем пункте.

- $S_j \sim IW(k_S^2 \times p_{S_j} \times \widehat{\text{Var}}(A_{j,OLS}), p_{S_j}), j = 1, \dots, n - 1$

Здесь значение  $k_S$  одно и то же для всех блоков матрицы  $S$ , равное 0.1. А вот количество степеней свободы выбирается в зависимости от размерности соответствующего блока матрицы  $S$  —  $p_{S_j} = j + 1$ .

Заметим, что математическое ожидание случайной матрицы, имеющей обратное распределение Уишарта  $IW(C, b)$ , равно  $\frac{C}{b-q-1}$ , где  $q$  - размерность матрицы  $C$ . И при  $b \rightarrow \infty$  математическое ожидание стремится к 0. Таким образом, чем большее число степеней свободы задается в априорном распределении гиперпараметров, тем меньшую волатильность имеют шоки в процессах случайного блуждания. Если же говорить о выбранных значениях  $k_Q, k_S, k_W$ , то автор утверждает, что они позволяют получить неинформативное априорное распределение, гибко подстраивающееся под данные модели.

### 2.4.3. Алгоритм МСМС

К сожалению, знать априорные распределения параметров и функцию правдоподобия недостаточно для того, чтобы получить апостериорное рас-

пределение. Как видно из формулы (12), это позволит нам получить только функцию, пропорциональную апостериорному распределению. В теории не составляет труда найти значение недостающей константы по формуле полной вероятности  $P(Y_t) = \int P(Y_T|\Theta) P(\Theta)d\Theta$ , но на практике очень часто посчитать такой многомерный интеграл не представляется возможным. Существует два способа решения этой проблемы: использование сопряженных распределений и применение алгоритмов МСМС для получения выборки из апостериорного распределения.

Сопряжённое априорное распределение подбирается таким образом, чтобы при вычислении его произведения с функцией правдоподобия, получалось апостериорное распределение того же семейства. В таком случае апостериорное распределение выписывается в явном виде.

Второй способ получить апостериорное распределение — использование методов МСМС (Markov Chain Monte Carlo). Данные методы помогают получить выборку из апостериорного распределения, которая в дальнейшем может быть использована для получения интересующих нас характеристик оцениваемых параметров. Существует три основных вида алгоритмов, позволяющих получить выборку из апостериорного распределения — сэмплирование по Гиббсу<sup>2</sup>, алгоритм Метрополиса-Гастингса<sup>3</sup> и гибридный метод Монте-Карло<sup>4</sup>. В основе всех этих методов лежит следующий принцип: строится цепь Маркова с заданными вероятностями перехода из одного состояния в другое. Стационарное состояние такой цепи и является плотностью искомого апостериорного распределения. Алгоритм иницируется в какой-то точке распределения и дальше, используя определенные критерии, начинает двигаться из одной точки в другую. При доста-

---

<sup>2</sup>Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6), 721-741.

<sup>3</sup>Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1), 97-109.

<sup>4</sup>Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. (1987). Hybrid monte carlo. *Physics letters B*, 195(2), 216-222.

точно большом количестве шагов полученная выборка будет достаточно близка к выборке из искомого апостериорного распределения.

Для получения выборки из апостериорного распределения в модели TVP-BVAR используется схема Гиббса. Рассмотрим алгоритм из статьи Primiceri, Del Negro (2015). Для простоты обозначим  $B^T$  все коэффициенты  $\{B_t\}_{t=1}^T$ , аналогичные обозначения введем для  $\Omega^T$  и  $A^T$  и  $y^T$ . А все ковариационные матрицы шоков объединим в матрицу  $V$ . В таком случае, алгоритм формулируется следующим образом:

**Шаг 1** Инициализируем алгоритм, задав значения параметров  $A^T, \Omega^T, V, s^T$ ;

**Шаг 2** Генерируем значения  $B^T$  из распределения  $P(B^T | A^T, \Omega^T, V, y^T)$ ;

**Шаг 3** Генерируем элементы  $Q$  из распределения  $P(Q | A^T, B^T, \Omega^T, y^T)$ , являющегося обратным распределением Уишарта;

**Шаг 4** Генерируем элементы  $A^T$  из распределения  $P(A^T | B^T, \Omega^T, V, y^T)$ ;

**Шаг 5** Генерируем элементы блоков  $S$  из распределения  $P(S | A^T, B^T, \Omega^T, y^T)$ , являющегося обратным распределением Уишарта;

**Шаг 6** Генерируем вспомогательную дискретную переменную  $s^T$  из распределения  $P(s^T | A^T, B^T, \Omega^T, V, y^T)$ ;

**Шаг 7** Генерируем элементы  $\Omega^T$  из распределения  $P(\Omega^T | A^T, B^T, V, y^T, s^T)$ ;

**Шаг 8** Генерируем элементы  $W$  из распределения  $P(W | A^T, B^T, \Omega^T, y^T)$ , являющегося обратным распределением Уишарта;

**Шаг 8** Возвращаемся к пункту 2.

На шаге 2, 4 и 7 для получения реализации случайных величин  $B^T, A^T$  и  $\Omega^T$  используется алгоритм из статьи Carter, Kohn (1994), в которой было показано как можно сгенерировать параметры в линейных моделях пространства состояний, используя фильтр Калмана.

Шаги 3, 5 и 8 представляют собой стандартную процедуру получения реализации случайной величины, имеющей обратное распределение Уишарта.

Применить алгоритм Carter, Kohn на шаге 7 напрямую, как это делается на 2 и 4 шаге, невозможно, так как если мы представим модель в линейной форме, как того требует данный алгоритм, то шоки в ней будут иметь не нормальное распределение, предполагаемое в модели, а представляют собой логарифм распределения хи-квадрат. Для решения этой проблемы используется алгоритм из работы Kim, Shephard and Chib (1998), где логарифм хи-квадрат распределения аппроксимируется набором из 7 нормальных распределений. Вспомогательная переменная  $s^T$  определяет, какой состав нормальных распределений будет использоваться в каждый момент времени для аппроксимации.

#### 2.4.4. Прогнозирование

Если основной целью является построение прогнозов, то помимо апостериорного распределения параметров необходимо получить апостериорное распределение будущих значений зависимых переменных при условии всех доступных данных:

$$\begin{aligned} P(Y_{T+1:T+H}|Y_T) &= \frac{P(Y_{T+1:T+H}, Y_T)}{P(Y_T)} = \\ &= \frac{\int P(Y_{T+1:T+H}|Y_T, \Theta) \cdot P(Y_T|\Theta) \cdot P(\Theta)d\Theta}{\int P(Y_T|\Theta) \cdot P(\Theta)d\Theta} \end{aligned} \quad (13)$$

Получение предиктивного распределения на практике осуществляется не напрямую с использованием выражения (13), а через апостериорное распределение параметров.

$$\begin{aligned} P(Y_{T+1:T+H}|Y_T) &= \int P(Y_{T+1:T+H}|Y_T, \Theta) \cdot P(\Theta|Y_T) d\Theta = \\ &= \int \prod_{t=T+1}^{T+H} P(Y_t|Y_T, \Theta) \cdot P(\Theta|Y_T) d\Theta \end{aligned}$$

Как и в случае с апостериорным распределением посчитать предиктивное распределение аналитически часто бывает невозможно, поэтому для получения прогнозов также используются МСМС алгоритмы. Рассмотрим алгоритм для получения выборки из предиктивного распределения, использующийся в пакете `bvarsv`.

**Шаг 1** Генерируем параметры из апостериорного распределения на  $H$  периодов вперед;

**Шаг 2** Генерируем вектор ошибок  $\varepsilon_{T+j}, j = 1, \dots, H$  из стандартного нормального распределения;

**Шаг 3** Используя значения, полученные на предыдущих шагах, генерируем будущие значения переменных, используя уравнение (6);

**Шаг 4** Для каждого  $j = 1, \dots, H$  считаем математическое ожидание и ковариационную матрицу будущего значения  $Y_{T+j}$ .

Чтобы понять, что именно считать прогнозом модели (это может быть, например, математическое ожидание предиктивного распределения, его мода или медиана), нужно определить функцию потерь, которая зависит от будущих значений переменной и наших прогнозов. Определившись с функцией потерь, мы можем найти такую функцию для определения прогноза, при которой значение функции потерь минимально. Чаще всего используется функция потерь, представляющая собой квадрат отклонений прогноза от реального значения переменной  $Y$ . В таком случае оптимальной



функцией для определения прогноза является условное математическое ожидание  $E(Y_{T+1:T+H}|Y_T)$ . В данной работе в качестве точечного прогноза также используется математическое ожидание.

## 3. Практическая часть

### 3.1. Данные

Рассмотрим данные, которые используются в данной работе для построения прогнозов. Всего было взято 33 временных ряда, список которых подробно представлен в Приложении в таблицах 6 и 7. Ряды условно делятся на 2 группы — макроэкономические и финансовые. Макроэкономический набор данных состоит из 9 временных рядов. Каждый временной ряд представляет собой ежемесячные наблюдения с января 2000 года по январь 2017 года включительно. Из 204 имеющихся наблюдений 180 вошли в тренировочную выборку, а оставшиеся 24 использовались для проверки качества полученных прогнозов.

Для набора были выбраны показатели, оказывающее непосредственное влияние на российскую экономику. В основном данный набор состоит из макроэкономических показателей, таких как индекс промышленного производства, инфляция (индекс потребительских цен), денежный агрегат M2, межбанковская процентная ставка, безработица и номинальная заработная плата; помимо этого есть важные финансовые показатели: индекс РТС, характеризующий состояние финансовых рынков страны, курс доллара, а также цена на нефть, которая, в силу высокой зависимости экономики страны от ее экспорта, может оказывать достаточно сильное влияние на макроэкономические показатели<sup>5</sup>. На рисунке 1 показаны все ряды, включенные в макроэкономический набор.

Прежде чем строить модели с использованием описанных выше данных, они были подвержены некоторым преобразованиям. Для начала был проведен тест на единичный корень и ряды, для которых гипотеза об отсутствии единичного корня отвергалась (процедура теста подробно описана

---

<sup>5</sup>например, данный вопрос подробно обсуждается в статье Kalcheva, Oomes "Diagnosing Dutch Disease: Does Russia Have the Symptoms?"

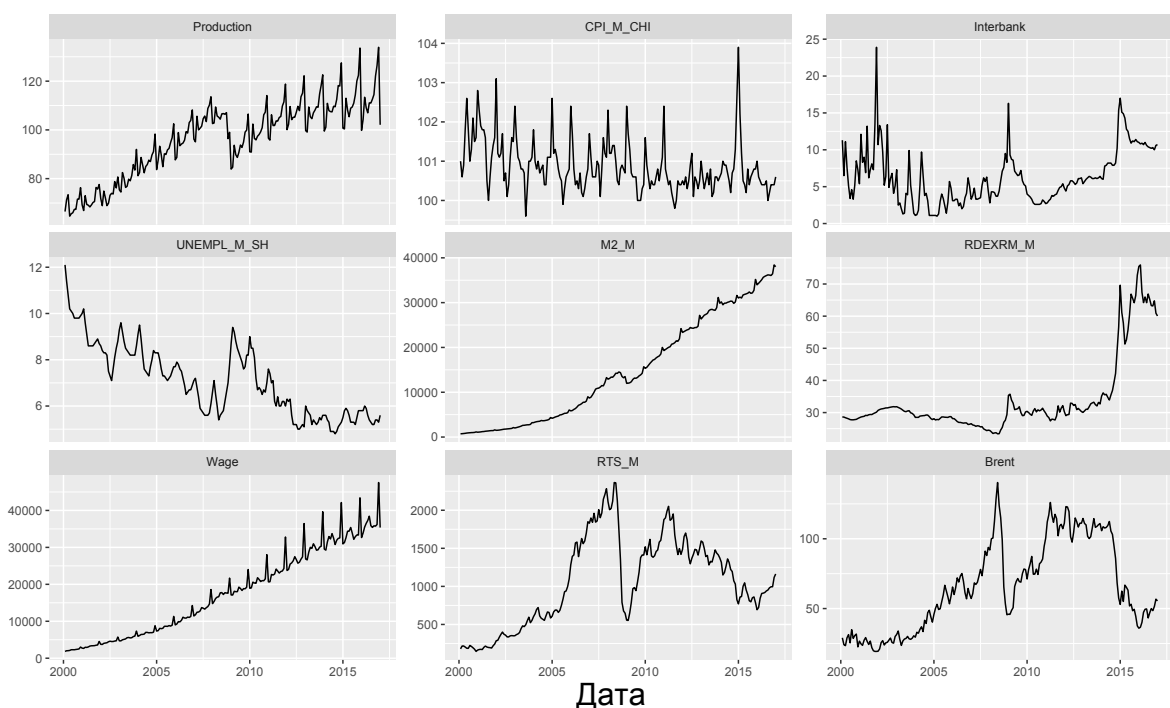


Рис. 1: Макроэкономические данные до преобразований

в разделе 2.2), были заменены на их первую разность. В таблице 6 показан порядок разности, использовавшийся для каждого из рядов. Заметим, что в результате проверки рядов на стационарность, мы везде перешли от абсолютных значений к приросту показателя и в дальнейшем работали именно с приростами. Далее данные были приведены к единому масштабу посредством следующего преобразования:

$$\tilde{y}_t = \frac{y_t - \bar{y}}{sd(y)} \quad (14)$$

Это было сделано для того, чтобы упростить вычисления при построении многомерных моделей. Прогнозы, полученные на отмасштабированных данных, возвращались к исходному масштабу для последующего расчета ошибки прогноза. На рисунке 2 представлены ряды из макроэкономического набора после всех преобразований.

Второй набор данных, который условно был назван финансовым, состоит из 24 временных рядов, каждый из которых представляет собой дневные цены акций крупнейших российских компаний, торгующихся на Мос-

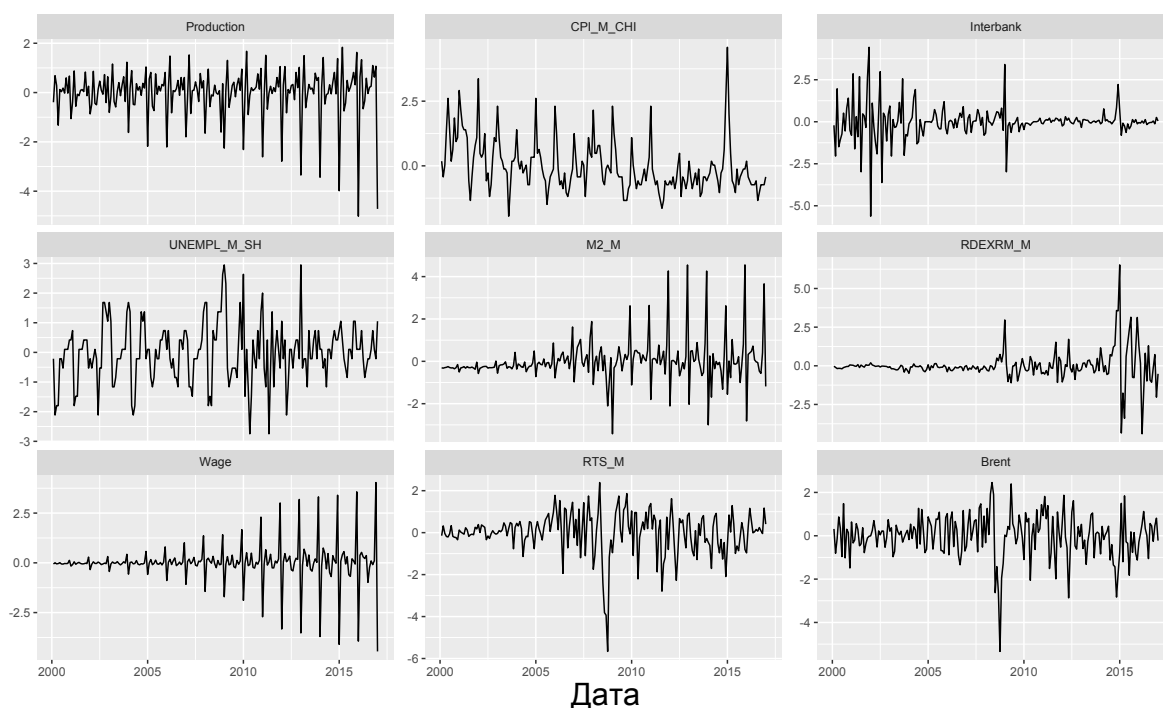


Рис. 2: Макроэкономические данные после преобразований

ковской бирже. В выборку включены данные за период с 7 марта 2014 по 20 апреля 2017 восьми различных индустрий: банки и финансы, машиностроение, металлы и добыча, нефть и газ, потребительские товары, телекоммуникации, химическое производство и электроэнергетика. Внутри каждой индустрии было выбрано 3 компании с наибольшей рыночной капитализацией, торгующиеся на рынке не менее трех лет. Второе условие было введено, чтобы иметь достаточное число наблюдений для построения моделей. В таблице 7 представлен полный список выбранных компаний. Из 815 доступных наблюдений, 555 наблюдений вошли в тренировочную выборку, а 260 — в тестовую.

Как и в случае с макроэкономическими данными, при построении моделей использовались модифицированные ряды. Во-первых, все цены были преобразованы в приросты цен, то есть была взята первая разность. Во-вторых, данные были приведены к единому масштабу при помощи преобразования (14). Итоговые прогнозы, качество которых сравнивалось для различных моделей, представляли собой приросты цен, которым был воз-

вращён исходный масштаб. На рисунке 3 показаны временные ряды из отрасли банков и финансов до и после трансформации.

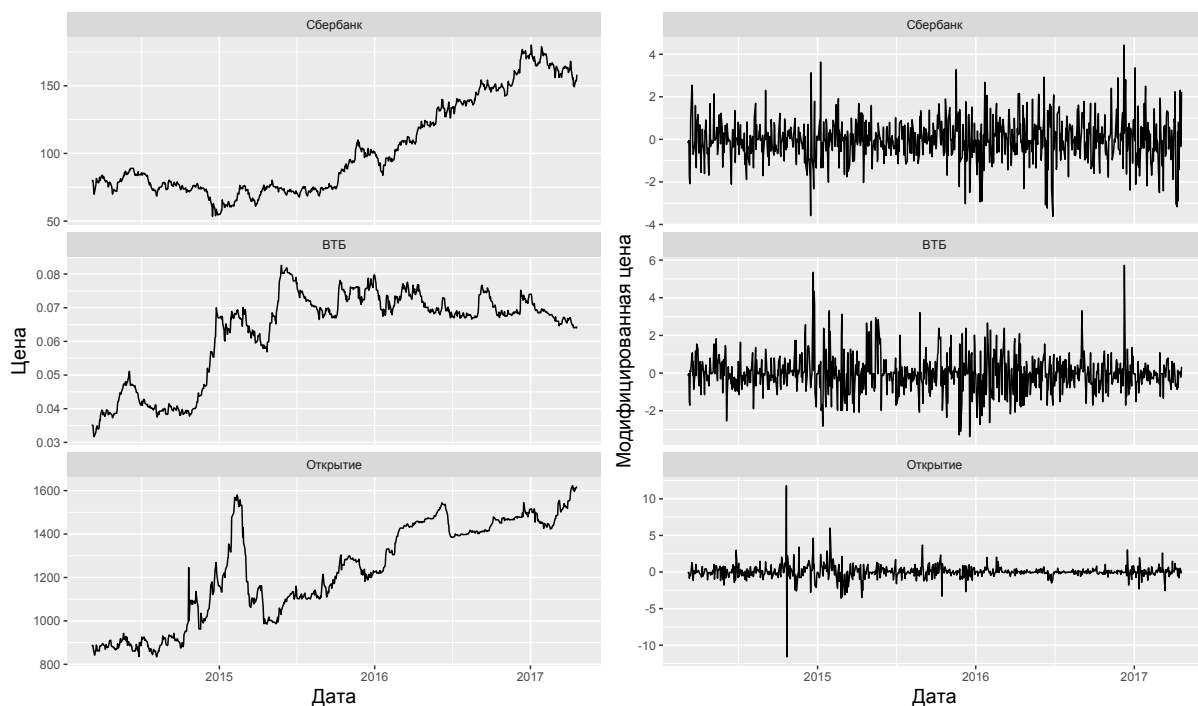


Рис. 3: Акции индустрии банки и финансы до и после преобразований

## 3.2. Методология

В данном разделе будет описан процесс построения моделей: выбор спецификаций, оценка коэффициентов, получение прогнозов, а также методы сравнения качества полученных прогнозов, которые использовались в данной работе. Заметим также, что методология для макроэкономических данных и для акций будет иметь некоторые различия.

### 3.2.1. Построение прогноза одномерных моделей

Одномерные модели: ARIMA, экспоненциальное сглаживание и тета-метод, строятся для всех типов данных. Для этого используются функции пакета `forecast`, предоставляющие возможность автоматического выбора оптимальной спецификации. Они используют алгоритмы, описанные в разделах 2.1-2.3 данной работы.

Далее строится гибридный прогноз. В этих целях для каждого ряда определяются те одномерные модели, которые имеют меньшую среднеквадратичную ошибку на тестовой выборке чем модель случайного блуждания. Однако, оказалось, что для нескольких временных рядов такая модель только одна. В таких случаях было решено строить гибридный прогноз по всем трем одномерным моделям.

Существует несколько способов получения гибридного прогноза. Самый простой — использовать среднее арифметическое имеющихся прогнозов. Даже такой простой шаг помогает в среднем улучшить качество прогнозов <sup>6</sup>. Более ресурсоемкий (в вычислительном плане), но при этом показавший лучшие результаты, подход — использование кросс-валидации для расчета ошибки модели на тренировочной выборке. В работе использовался именно этот подход, поэтому рассмотрим подробнее принцип его работы.

Так как мы имеем дело с временными рядами, у нас нет возможности исключать наблюдения случайным образом и строить модель на оставшихся данных. Поэтому кросс-валидация для временных рядов подразумевает использование скользящего окна. Принцип работы данного метода заключается в следующем: выбирается фиксированное число наблюдений  $n < N$ , где  $N$  — размер тренировочной выборки, а также горизонт прогноза  $h$ . Используя  $n$  первых наблюдений оценивается интересующая нас модель и строится прогноз на  $h$  периодов вперед. Затем, мы «сдвигаемся» на  $h$  периодов вперед, оцениваем модель на наблюдениях  $h : h + n$  и снова строим прогноз. Процедура повторяется до тех пор, пока мы не получим прогнозы для всех наблюдений тестовой выборки, кроме первых  $n$ .

Используя полученные значения, считается ошибка прогноза. Например, его среднеквадратичное отклонение от действительного значения переменной (RMSE). Получив ошибку на тренировочной выборке, для каж-

---

<sup>6</sup>см. статью Bates and Granger, The Combination of Forecasts (1969)

дой модели, используемой для построения гибридного прогноза, считаются веса по следующей формуле:

$$w_i = \frac{\frac{1}{RMSE_i}}{\sum_{j=1}^k \frac{1}{RMSE_j}}, \quad (15)$$

где  $k$  — общее число моделей. Заметим, что чем больше ошибка на тестовой выборке, тем меньший вес будет иметь модель при построении гибридного прогноза. Полученные из уравнения (15) веса присваиваются прогнозам разных моделей и таким образом получается гибридный прогноз.

В результате, применяя описанную выше методологию, для каждого временного ряда мы рассчитываем 4 разных прогноза с использованием одномерных моделей.

### 3.2.2. Построение прогноза TVP-BVAR

В первую очередь, в рамках каждого из двух наборов данных были выбраны группы переменных, которые впоследствии использовались для построения байсовских векторных авторегрессий с переменными параметрами. В связи с ограничениями пакета `bvarsv`, с помощью которого реализовано построение моделей TVP-BVAR, оценить параметры векторной авторегрессии можно только для ограниченного числа переменных и лагов. Как уже упоминалось выше, включение каждой последующей переменной или увеличение числа лагов приводит к значительному росту количества оцениваемых параметров, что становится причиной возникновению ошибок в алгоритме семплирования. Именно поэтому в рамках данного исследования не было возможности построить многомерную модель, включающую в себя все временных ряды из набора данных.

Важным критерием для определения таких групп является «схожесть» переменных между собой. Для векторной авторегрессии большую роль играет совместная динамика переменных, поэтому нет смысла включать

в модель переменные, которые совершенно не связаны между собой. Руководствуясь данным предположением, для финансового набора данных было решено строить отдельные модели для каждой индустрии. Логично предположить, что крупнейшие компании одной индустрии могут быть подвержены одним и тем же внешним шокам, более того, шоки, влияющие на одну компанию могут влиять на ожидания игроков финансового рынка относительно других компаний данной индустрии и, следовательно, приводить к колебаниям цен на акции всех компаний. Таким образом, было построено 8 TVP-BVAR моделей для финансового набора данных, в каждом из которых использовалось по 4 лага.

В случае с макроэкономическими данными гораздо сложнее выделить группы наиболее тесно связанных переменных. Поэтому в работе использовался метод поиска схожих объектов (переменных) путем снижения размерности данных — t-SNE (t-distributed Stochastic Neighbour Embedding). Для поиска оптимального пространства малой размерности в данном методе сначала определяется коэффициенты схожести всех точек исходного пространства —  $p_{i,j}$ . Этот коэффициент зависит от расстояния между этими точками  $D_{i,j}$  и считается следующим образом:

$$p_{j|i} = \frac{\exp(-\|D_{i,j}\|^2/2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|D_{i,k}\|^2/2\sigma_i^2)}$$

$$p_{i,j} = \frac{p_{j|i} + p_{i|j}}{2n}$$

Параметр  $\sigma$  выбирается таким образом, чтобы перплексивность случайной величины  $p_{i|j}$  была близка к указанному пользователем значению. Фактически, параметр перплексивности контролирует эффективное число «соседей» точки  $i$ , которые имеют значимый вес при переходе в пространство меньшей размерности. В пространстве низкой размерности расстояния между точками имеют распределение Стюдента с 1 степенью свободы



(отсюда и название метода) и рассчитываются следующим образом:

$$q_{i,j} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|y_k - y_l\|^2)^{-1}}$$

Положение точек в новом пространстве определяется так, чтобы получить как можно более схожие значения  $q_{i,j}$  и  $p_{i,j}$ .

На рисунке 4 представлены результаты применения данного алгоритма к макроэкономическим данным.

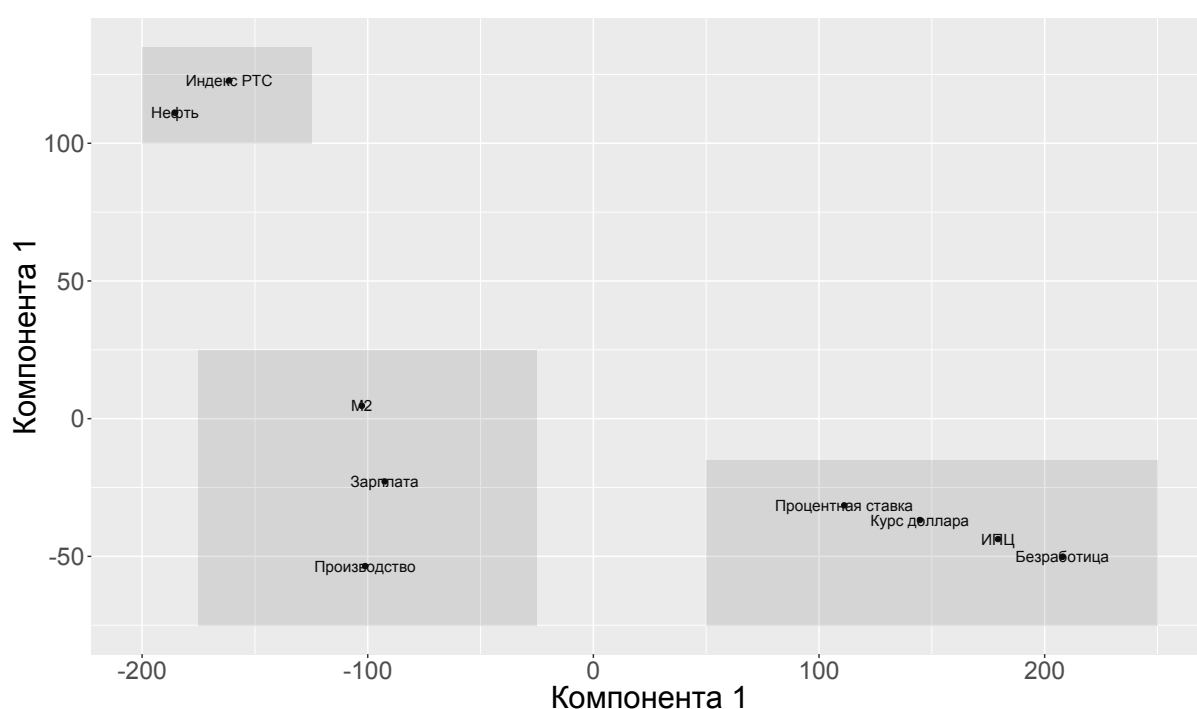


Рис. 4: Макроэкономические данные, деление на кластеры

В результате снижения размерности данных до двух наблюдений мы получили три явных кластера. Построим 6 разных моделей - для каждого кластера отдельно, а также для всех комбинаций двух кластеров между собой. Количество лагов для каждой модели выбиралось с учетом ограничений статистического пакета `bvarsv`. Так, для больших моделей, в которые вошли данные двух кластеров, использование более чем одного лага приводило к возникновению ошибок. В таблице 2 представлены спецификации всех построенных в результате моделей.

Таблица 2: TVP-BVAR модели для макроэкономических рядов

| Переменные                          | Количество лагов | Название |
|-------------------------------------|------------------|----------|
| y, $\pi$ , r, u, m, $e_{doll}$ , w  | 1                | TVP-B1   |
| y, m, w, rts, oil                   | 1                | TVP-B2   |
| $\pi$ , r, u, $e_{doll}$ , rts, oil | 1                | TVP-B3   |
| y, m, w                             | 6                | TVP-S6   |
|                                     | 3                | TVP-S3   |
|                                     | 12               | TVP-S12  |
| rts, oil                            | 6                | TVP-S6   |
|                                     | 3                | TVP-S3   |
|                                     | 3                | TVP-S3   |
| $\pi$ , r, u, $e_{doll}$            | 3                | TVP-S3   |
|                                     | 1                | TVP-S1   |

Источник: Расчеты автора

При построении моделей с использованием байесовского подхода важным этапом является выбор количества итераций, так как от этого сильно зависит качество полученных результатов. На первых шагах алгоритма семплирования выборка достаточно далека от истинного апостериорного распределения и количество итераций должно быть достаточно большим, чтобы полученная марковская цепь «сошлась» к своему стационарному состоянию. Для того, чтобы не получать искаженные оценки, в результирующую выборку не включают первую часть полученных случайных значений. Считается, что эти итерации служат для «прожига» (burn - in) Марковского процесса. К сожалению, нельзя точно проверить или оценить степень сходимости цепи, однако есть критерии, которые помогают идентифицировать отсутствие сходимости, основной из которых — наличие высокой и устойчивой автокорреляции в полученной выборке. Для снижения вероятности получения цепи с высокой автокорреляцией часто не только исключают первые итерации, но и «прореживают» оставшиеся элементы, выбирая только каждый второй, пятый или даже десятый элемент.

При построении модели TVP-BVAR генерируется марковская цепь с очень большим числом звеньев (значений). Так, в модели, построенной для макроэкономических данных с пятью переменными и 1 лагом (TVP-B2), количество сгенерированных выборок итоговых коэффициентов модели и элементов ковариационных матриц остатков составляет 7645, то есть 55 выборок для каждого из 139 моментов времени, 25 из которых представляют собой элементы ковариационной матрицы остатков, а 30 — коэффициенты модели (5 уравнений с 5 объясняющими переменными и одним свободным коэффициентом в каждом). В результате, проверка сходимости каждой цепи связано с очень высокими трудозатратами, поэтому было принято решение проверить сходимость 4170 цепей в описанной выше модели, с использованием различного количества итераций.

В результате было обнаружено, что при использовании 5000 итераций для «розжига» и включения в итоговую выборку каждого 10-го элемента цепи из 20000 случайно сгенерированных элементов, автокорреляция 20-го порядка в каждой из более чем четырех тысяч цепей крайне редко превосходит значение 0.1, что показано на рисунке 5. Опираясь на данные результаты, последующие байесовские векторные авторегрессии оценивались с таким же количеством итераций.

### **3.2.3. Сравнение прогнозов**

Прогнозы, полученные в результате, сравниваются на основании отношения среднеквадратичной ошибки прогноза данной модели к среднеквадратичной ошибке прогноза случайного блуждания. То есть ошибка  $i$ -ой модели в прогнозировании  $k$ -ой переменной на горизонте  $h$  рассчитывается следующим образом:

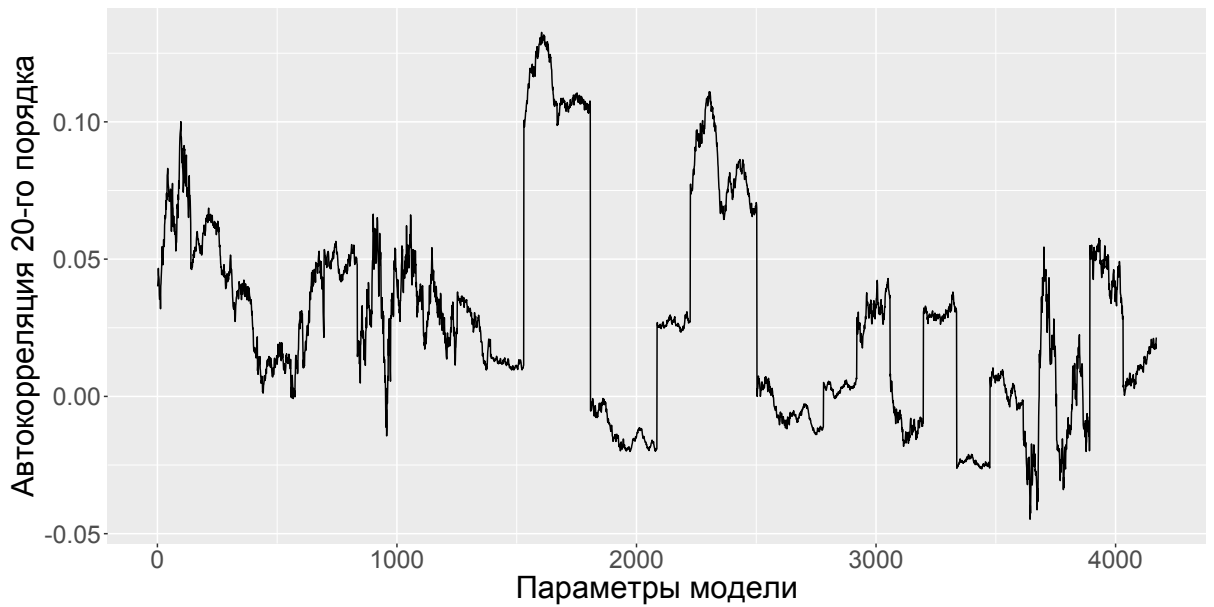


Рис. 5: Автокорреляция для параметров модели TVP-B2

$$Error_{k,h}^{(i)} = \frac{RMSE_{k,h}^{(i)}}{RMSE_{k,h}^{(RW)}}, \quad \text{где}$$

$$RMSE_{k,h}^{(i)} = \sqrt{\frac{1}{h} \sum_{j=1}^h (y_{T+j} - \hat{y}_{T+j})^2}$$

Модель случайного блуждания предполагает, что лучший прогноз всех будущих значений ряда — это последнее известное наблюдения. Такой прогноз еще называют наивным. В следующей части работы будут представлены полученные результаты оценки моделей и проведено сравнение качества прогнозов.

### 3.3. Результаты

В общей сложности было оценено 159 моделей. Для макроэкономических данных было построено по 4 одномерные модели для каждого из 9 рядов, а также 10 моделей TVP-BVAR, описанных в таблице 2. Для сравнения прогнозов на разных горизонтах была рассчитана относительная ошибка

прогноза на 3, 6, 12 и 24 периода вперед. Для финансовых рядов было оценено 96 одномерных моделей, 4 модели на каждый из 24 временных рядов, а также 8 моделей байесовской векторной авторегрессии, отдельно для каждой индустрии. Так как финансовые данные содержат ежедневные наблюдения, то горизонты прогноза, для которых рассчитывалась относительная ошибка равнялись 7, 14, 30, 60 и 260 наблюдениям.

### 3.3.1. Результаты прогнозирования макроэкономических данных

На рисунке 6 представлена динамика ошибок прогноза всех макроэкономических переменных для разных горизонтов. Заметим, что каждая из переменных участвовала в двух «больших» моделях TVP-BVAR (то есть содержащих более 4 переменных) и, как минимум, в двух «маленьких» (содержащих 4 и менее переменных). На графике представлены относительные ошибки только двух моделей: лучшей из двух «больших» моделей и лучшей из двух или трех «маленьких». Критерием выбора служила средняя ошибка по 4 горизонтам.

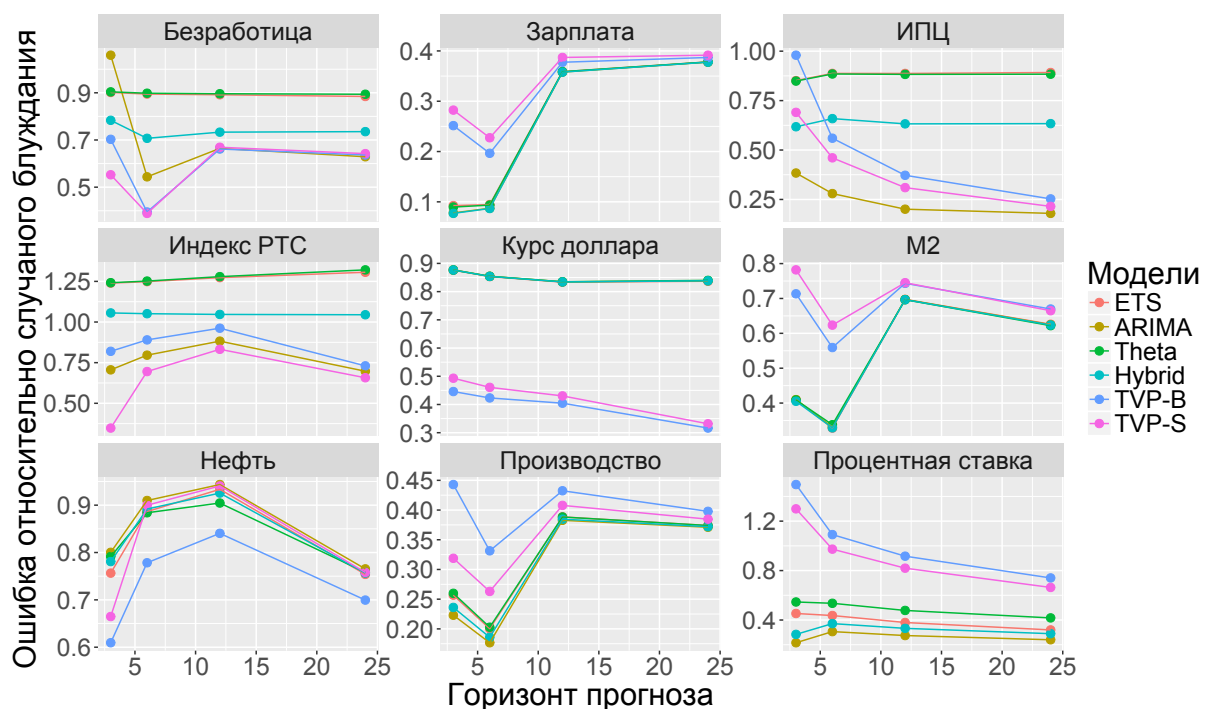


Рис. 6: Относительные ошибки прогноза для макроэкономических данных

На графиках видно, что TVP-BVAR оказался наиболее эффективным при прогнозировании цены на нефть и курса доллара к рублю. Причем, если в первом случае хорошие результаты по сравнению с другими моделями показала только модель с большим количеством переменных, то курс доллара одинаково хорошо был предсказан обеими спецификациями. Заметим, что не наблюдается значительного преимущества использования гибридного прогноза по сравнению с обычными одномерными моделями.

График дает общее представление о том, как соотносятся ошибки прогнозов различных моделей. Поэтому, чтобы более формально сравнить модели между собой, рассмотрим таблицу 3.

Таблица 3: Результаты прогнозирования макроэкономических показателей

|                   | 3 месяца | 6 месяцев | 12 месяцев | 24 месяца |
|-------------------|----------|-----------|------------|-----------|
| Производство      | ARIMA    | ARIMA     | ARIMA      | ARIMA     |
| ИПЦ               | ARIMA    | ARIMA     | ARIMA      | ARIMA     |
| Процентная ставка | ARIMA    | ARIMA     | ARIMA      | ARIMA     |
| Безработица       | TVP-S1   | TVP-S1    | TVP-B3     | ARIMA     |
| M2                | TVP-B1   | ETS       | TVP-B1     | Theta     |
| Курс доллара      | TVP-B1   | TVP-B1    | TVP-B1     | TVP-B3    |
| Зарплата          | Hybrid   | Hybrid    | TVP-B1     | Hybrid    |
| Индекс РТС        | TVP-S6   | TVP-S6    | TVP-S6     | TVP-S6    |
| Нефть             | TVP-B2   | TVP-B2    | TVP-B2     | TVP-B2    |

Источник: расчеты автора

Здесь представлены модели, которые имеют наименьшую относительную ошибку для данной переменной и горизонта прогноза. Не сложно заметить, что среди всех одномерных моделей явным лидером стала ARIMA. Гибридные прогнозы показали высокий результат только для номинальной

заработной платы. В то время как модели TVP-BVAR предоставили лучший прогноз ровно в половине случаев. Также заметим, что при прогнозировании на длительный период времени, «лучшими» чаще оказываются одномерные модели.

Так для каждой переменной из данного набора было построено несколько различных TVP-BVAR моделей, мы можем посмотреть, как соотносятся между собой ошибки прогноза моделей с разным количеством переменных и лагов. В таблице 4 представлены модели с наименьшими ошибками, но только среди многомерных моделей. Из таблицы видно, что «большие» модели, в которые включалось более четырех переменных, чаще оказываются среди лучших. Еще одно важное наблюдение заключается в том, что подавляющая часть «маленьких» моделей в таблице была построена с 3 или 1 лагом. Модели с 6 лагами представлены в небольшом количестве, а модели с 12 лагами ни разу не показали лучших результатов.

Таблица 4: Сравнение разных спецификаций TVP-BVAR

|                   | 3 месяца | 6<br>месяцев | 12<br>меся-<br>цев | 24<br>меся-<br>ца |
|-------------------|----------|--------------|--------------------|-------------------|
| Производство      | TVP-S3   | TVP-S3       | TVP-S3             | TVP-S3            |
| ИПЦ               | TVP-B1   | TVP-B1       | TVP-B1             | TVP-S1            |
| Процентная ставка | TVP-B1   | TVP-S1       | TVP-S1             | TVP-S1            |
| Безработица       | TVP-S1   | TVP-S1       | TVP-B3             | TVP-B3            |
| M2                | TVP-B1   | TVP-B1       | TVP-B1             | TVP-S3            |
| Курс доллара      | TVP-B1   | TVP-B1       | TVP-B1             | TVP-B3            |
| Зарплата          | TVP-S6   | TVP-B2       | TVP-B1             | TVP-B2            |
| Индекс РТС        | TVP-S6   | TVP-S6       | TVP-S6             | TVP-S6            |
| Нефть             | TVP-B2   | TVP-B2       | TVP-B2             | TVP-B2            |

Источник: расчеты автора

### 3.3.2. Результаты прогнозирования финансовых данных

Перейдем к данным из финансовой выборки. Напомним, что всего прогнозировалось 24 временных ряда, представляющих изменение цен акций крупнейших российских компаний из 8 индустрий. На рисунке 7 представлены относительные ошибки прогноза для каждой из 24 переменных, отсортированные по отраслям. В данном случае каждая переменная участвовала только в одной TVP-BVAR модели, построенной для трех компаний одной отрасли. На графике можно видеть компании, изменение цен которых лучше всего спрогнозировала именно байесовская векторная авторегрессия с переменными коэффициентами, например, Лукойл, Газпром, М.Видео (для небольших горизонтов). А для таких компаний как Магнит, Мегафон и Ростелеком одномерные модели оказались более эффективными. Однако, в целом, глядя на график, сложно определить фаворита. Явно видно, что во многих случаях разница в прогнозе достаточно маленькая.

Наконец, рассмотрим результаты, представленные в таблице 5. Как и в случае с графиком, изучая список моделей, имеющих наименьшую относительную ошибку, нельзя выявить модель, которая лучше остальных прогнозируют изменение цен акций. Тем не менее, стоит отметить, что более чем в трети случаев многомерная модель оказалась лучше одномерных. Также можно заметить, что высокие результаты TVP-BVAR для одной компании часто означают аналогичные результаты для других компаний той же отрасли. Такой результат кажется закономерным, так как модели строились для отрасли в целом.

Однако, из графика на рисунке 7 мы знаем, что для многих компаний разница в относительной ошибке крайне незначительна. Таким образом, не стоит делать вывод, что модели, представленные в таблице, действительно обеспечивают более высокое качество прогнозов. Дело в том, что если говорить об одномерных моделях, то использование замена одних



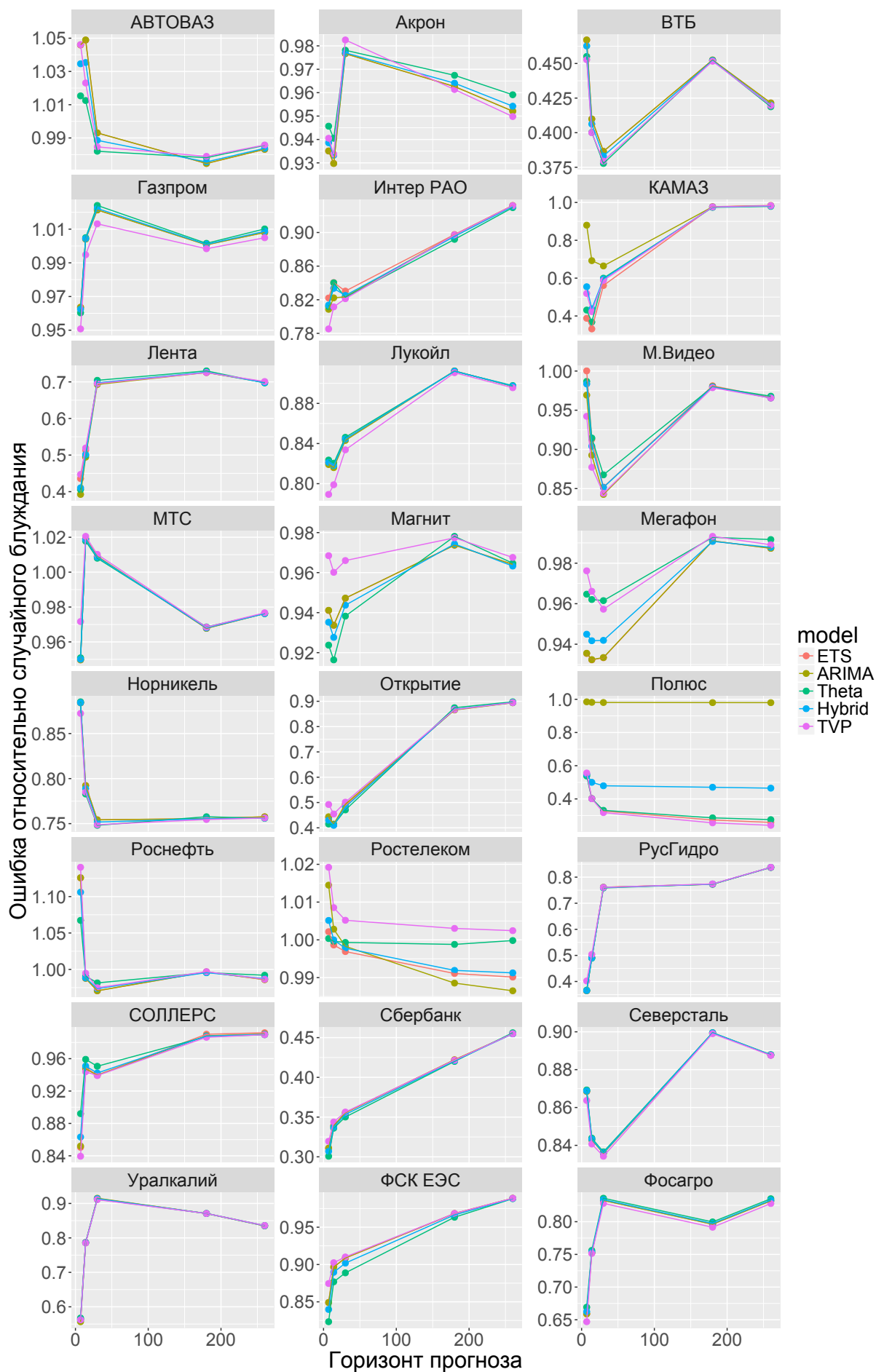


Рис. 7: Относительные ошибки прогноза финансовых данных

Таблица 5: Результаты прогнозирования финансовых данных

|            | h = 7 | h = 14 | h = 30 | h = 180 | h = 260 |
|------------|-------|--------|--------|---------|---------|
| Сбербанк   | Theta | Theta  | Theta  | Theta   | Hybrid  |
| ВТБ        | TVP   | TVP    | Theta  | TVP     | Theta   |
| Открытие   | Theta | Hybrid | Theta  | ETS     | ETS     |
| АВТОВАЗ    | Theta | Theta  | Theta  | ETS     | ETS     |
| КАМАЗ      | ARIMA | ARIMA  | ARIMA  | Theta   | Theta   |
| СОЛЛЕРС    | TVP   | TVP    | TVP    | TVP     | Theta   |
| Норникель  | TVP   | Theta  | Theta  | TVP     | Theta   |
| Полюс      | Theta | TVP    | TVP    | TVP     | TVP     |
| Северсталь | TVP   | TVP    | TVP    | TVP     | TVP     |
| Роснефть   | Theta | ETS    | ETS    | Theta   | ETS     |
| Газпром    | TVP   | TVP    | TVP    | TVP     | TVP     |
| Лукойл     | TVP   | TVP    | TVP    | TVP     | TVP     |
| Магнит     | Theta | Theta  | Theta  | ARIMA   | Hybrid  |
| Лента      | ETS   | ETS    | ETS    | TVP     | Theta   |
| М.Видео    | TVP   | TVP    | ETS    | TVP     | TVP     |
| МТС        | ARIMA | Theta  | Theta  | ARIMA   | Hybrid  |
| Мегафон    | ARIMA | ARIMA  | ARIMA  | Hybrid  | ARIMA   |
| Ростелеком | Theta | ARIMA  | ARIMA  | ETS     | ETS     |
| Уралкалий  | ETS   | TVP    | TVP    | ETS     | ETS     |
| Фосагро    | TVP   | TVP    | TVP    | TVP     | TVP     |
| Акрон      | ETS   | ETS    | ETS    | TVP     | TVP     |
| Интер РАО  | TVP   | TVP    | TVP    | Theta   | Theta   |
| РусГидро   | ARIMA | Theta  | Theta  | ARIMA   | ARIMA   |
| ФСК ЕЭС    | Theta | Theta  | Theta  | Theta   | Theta   |

Источник: расчеты автора

методов на другие даже для незначительного улучшения прогноза вполне оправдано, так как их оценка не требует большого количества вычислительных ресурсов. Если же рассматривать необходимость использования таких сложных методов как TVP-BVAR, стоит принимать во внимание, что оценка одной модели может занимать до нескольких часов или даже дней в зависимости от мощности используемого компьютера, а результаты часто незначительно превосходят результаты одномерных моделей. Таким образом, если стоит задача построение оперативного прогноза на ближайший период времени, стоит серьезно оценить вычислительные мощности используемого оборудования при построении таких моделей как TVP-BVAR.

## Заключение

В данной работе было проведено исследование прогнозов, построенных с помощью байесовской векторной авторегрессии с переменными параметрами и одномерных моделей, таких как ARIMA, экспоненциальное сглаживание, тета-метод. Целью работы было сравнить качество полученных прогнозов для различных типов данных и для горизонтов различной продолжительности. Для анализа было выбрано два набора данных: месячные макроэкономические показатели и дневные изменения цен акций крупнейших российских компаний.

После оценки моделей на тренировочной выборке были построены прогнозы и рассчитаны ошибки моделей по сравнению с моделью случайного блуждания. Мы увидели, что прогнозы TVP-BVAR оказались ближе всего к истинным значениям ровно в половине случаев для макроэкономических данных и более чем в трети случаев для финансового набора данных. Тем не менее, было замечено, что для финансовых данных разница в относительных ошибках незначительна, в результате чего нельзя однозначно говорить о преимуществе данной модели.

Серьезным ограничением при проведении данного исследования служили возможности используемого статистического пакета. Одним из возможных расширений данной работы могла бы служить оптимизация кода, которая позволила бы оценивать модель с большим количеством переменных и лагов. Предположительно, это позволило бы получить более точные прогнозы модели TVP-BVAR, так как на макроэкономических данных, где для каждой переменной было построено несколько многомерных моделей, «большие» модели оказались более эффективными.

# Приложения

## А. Макроэкономические данные

Таблица 6: Макроэкономические временные ряды

| Название   | Описание   | Единицы измерения   | Разность | Источник |
|------------|--|---------------------|----------|----------|
| y          | Индекс промышленного производства                                  | индекс (2010 = 100) | 1        | ФСГС     |
| $\pi$      | Индекс потребительских цен, изменение по отношению к прошлому году | %                   | 0        | ФСГС     |
| r          | Средняя за месяц межбанковская overnight ставка                    | %                   | 1        | ЦБ РФ    |
| u          | Удельная доля безработных среди экономически активного населения   | %                   | 1        | ФСГС     |
| m          | Денежный агрегат M2  | млрд. руб           | 1        | ЦБ РФ    |
| oil        | Стоимость одного баррел нефти                                      | долл                | 1        | Финам    |
| $e_{doll}$ | Номинальный обменный курс доллара                                  | руб за долл         | 1        | ММВБ     |
| w          | Среднемесячная номинальная заработная плата                        | руб                 | 1        | МЭР      |
| rts        | Композиторный индекс российского фондового рынка                   | пункты              | 1        | ЕАЭСД    |

ФСГС — [Федеральная служба государственной статистики](#)

ЦБ РФ — [Центральный банк Российской Федерации](#)

Финам — [Инвестиционная компания «ФИНАМ»](#)

ММВБ — [Московская биржа](#)

МЭР — [Министерство экономического развития](#)

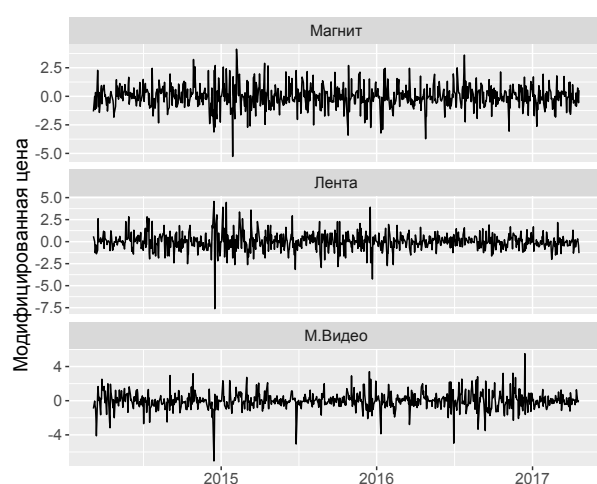
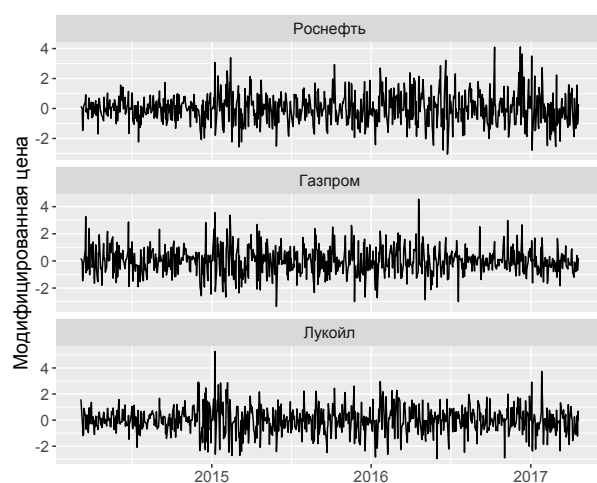
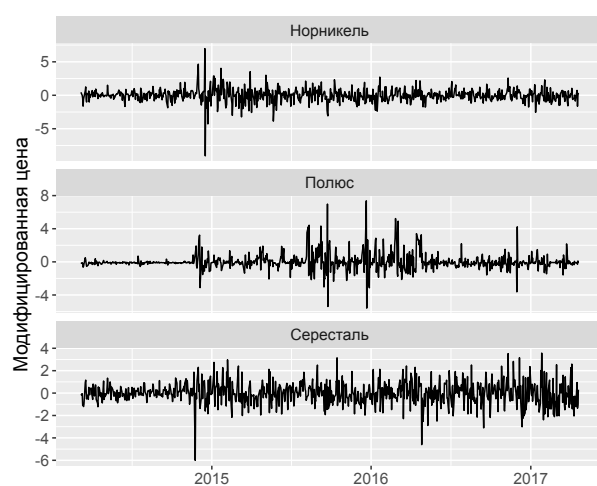
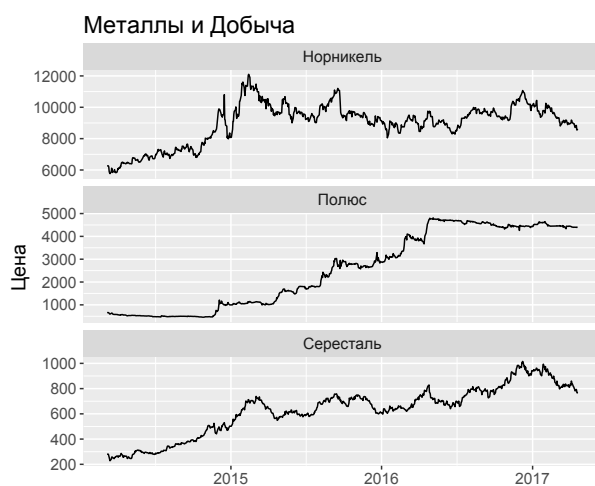
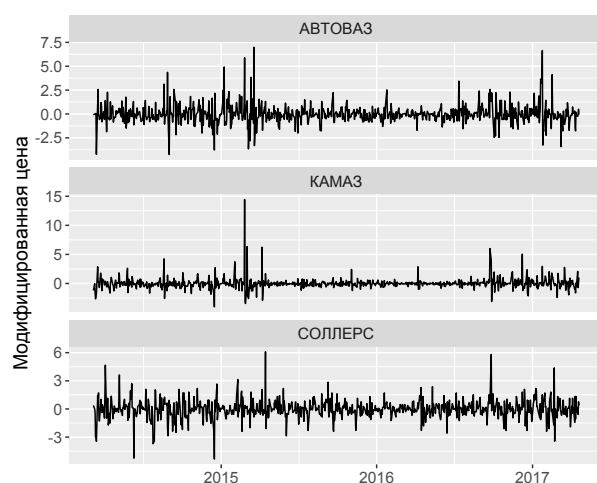
ЕАЭСД — [Единый архив экономических и социологических данных](#)

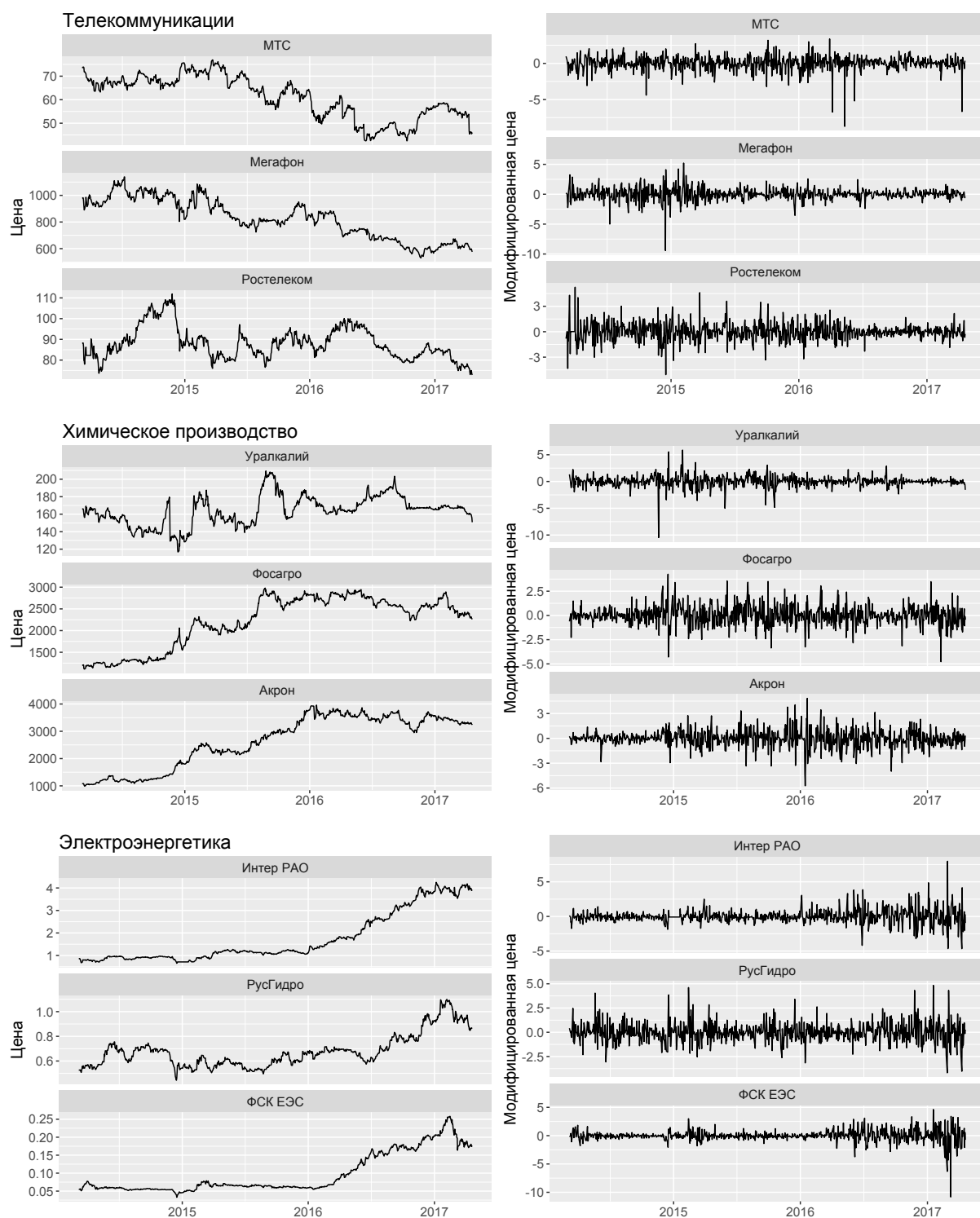
## В. Финансовые данные

Таблица 7: Финансовые временные ряды

| Индустрия                         | Название компании |
|-----------------------------------|-------------------|
| Банки и Финанасы                  | Сбербанк          |
|                                   | ВТБ               |
|                                   | Открытие          |
| Машиностроение                    | АВТОВАЗ           |
|                                   | КАМАЗ             |
|                                   | СОЛЛЕРС           |
| Металлы и Добыча                  | Норникель         |
|                                   | Полюс             |
|                                   | Северсталь        |
| Нефть и газ                       | Роснефть          |
|                                   | Газпром           |
|                                   | Лукойл            |
| Потребительские товары и торговля | Магнит            |
|                                   | Лента             |
|                                   | М.Видео           |
| Телекоммуникации, IT              | МТС               |
|                                   | Мегафон           |
|                                   | Ростелеком        |
| Химическое производство           | Уралкалий         |
|                                   | Фосагро           |
|                                   | Акрон             |
| Электроэнергетика                 | Интер РАО         |
|                                   | РусГидро          |
|                                   | ФСК ЕЭС           |

Источник данных: Терминал Thomson Reuters





## С. Расчеты

Все расчеты, проводившиеся в рамках данной работы, доступны в [репозитории на GitHub](#).



## Список литературы

- [1] Artis, M. J., & Zhang, W. (1990). BVAR forecasts for the G-7. *International journal of forecasting*, 6(3), 349-362.
- [2] Assimakopoulos, V., & Nikolopoulos, K. (2000). The theta model: a decomposition approach to forecasting. *International journal of forecasting*, 16(4), 521-530.
- [3] Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4), 451-468.
- [4] Beauchemin, K. R., & Zaman, S. (2011). A medium scale forecasting model for monetary policy.
- [5] Berg, T. O., & Henzel, S. (2013). Point and Density Forecasts for the Euro Area Using Many Predictors: Are Large BVARs Really Superior? (No. 155). ifo Working Paper.
- [6] Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: forecasting and control*, 1976. ISBN: 0-8162-1104-3.y.
- [7] Brown, R. G. (1959). *Statistical forecasting for inventory control*. McGraw/Hill.
- [8] Carriero, A., Kapetanios, G., & Marcellino, M. (2012). Forecasting government bond yields with large Bayesian vector autoregressions. *Journal of Banking & Finance*, 36(7), 2026-2047.
- [9] Carter, C. K., & Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 541-553.
- [10] Cogley, T., & Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Review of Economic dynamics*, 8(2), 262-302.

- [11] Del Negro, M., & Primiceri, G. E. (2015). Time varying structural vector autoregressions and monetary policy: a corrigendum. *The review of economic studies*, 82(4), 1342-1345.
- [12] Demeshev, B., & Malakhovskaya, O. (2015). Forecasting Russian macroeconomic indicators with BVAR.
- [13] Deutsch, M., Granger, C. W., & Teräsvirta, T. (1994). The combination of forecasts using changing weights. *International Journal of Forecasting*, 10(1), 47-57.
- [14] Doan, T., Litterman, R., & Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric reviews*, 3(1), 1-100.
- [15] Funke, M. (1990). Assessing the forecasting accuracy of monthly vector autoregressive models: The case of five OECD countries. *International Journal of Forecasting*, 6(3), 363-378.
- [16] Granger, C. W., & Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of forecasting*, 3(2), 197-204.
- [17] Holt, C. C. (1957). Forecasting seasonals and trends by exponentially weighted moving averages, *ONR Memorandum (Vol. 52)*, Pittsburgh, PA: Carnegie Institute of Technology. Available from the Engineering Library, University of Texas at Austin.
- [18] Hyndman, R. J., & Billah, B. (2003). Unmasking the Theta method. *International Journal of Forecasting*, 19(2), 287-290.
- [19] Hyndman, R. J., & Khandakar, Y. (2007). Automatic time series for forecasting: the forecast package for R (No. 6/07). Monash University, Department of Econometrics and Business Statistics.

- [20] Hyndman, R. J., Koehler, A. B., Snyder, R. D., and Grose, S. A state space framework for automatic forecasting using exponential smoothing methods // 2002 International Journal of Forecasting, 18, 439 - 454.
- [21] Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. The review of economic studies, 65(3), 361-393.
- [22] Koop, G., & Korobilis, D. (2010). Bayesian multivariate time series methods for empirical macroeconomics. Foundations and Trends® in Econometrics, 3(4), 267-358.
- [23] Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. Journal of econometrics, 54(1-3), 159-178.
- [24] Litterman, R. (1986). Forecasting with Bayesian vector autoregressions—Five years of experience: Robert B. Litterman, Journal of Business and Economic Statistics 4 (1986) 25–38. International Journal of Forecasting, 2(4), 497-498.
- [25] Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. Journal of the american statistical association, 55(290), 299-306.
- [26] Osborn, D. R. (1988). Seasonality and habit persistence in a life cycle model of consumption. Journal of Applied Econometrics, 3(4), 255-266.
- [27] Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. The Review of Economic Studies, 72(3), 821-852.
- [28] Sims, C. A. (1980). Macroeconomics and reality. Econometrica: Journal of the Econometric Society, 1-48.

- [29] Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management science*, 6(3), 324-342.
- [30] Yule, G. U. (1927). On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 226, 267-298.