

Утверждено приказом  
Министерства экономики  
Кыргызской Республики  
от 4 мая 2018 года №60

**Методическая основа  
прогнозирования потребности секторов  
экономики Кыргызской Республики в  
трудовых ресурсах**

# 1 Введение

Создание рабочих мест и хорошо функционирующий рынок труда - ключевые движущие факторы для сокращения бедности и развития. Рабочие места не только побочный продукт роста, но и сами по себе представляют цель развития и приоритет для политики. Часто представляя самый важный актив для бедных, рабочие места являются главным механизмом вывода людей из бедности. Таким образом, расширение возможностей трудоустройства и увеличение доходов для существующих рабочих мест представляют высший приоритет для разработчиков политики, как в развивающихся, так и в развитых странах. Это тем более верно в отношении таких стран, как Кыргызская Республика, в последние годы характеризующаяся растущим оттоком трудовых мигрантов.

В этом контексте крайне важно иметь возможность точно предсказывать развитие рынка труда на регулярной основе. Введение модели прогнозирования для кыргызского рынка труда является ключевым шагом в рамках более обширных продолжающихся усилий по укреплению потенциала, которые, на национальном уровне, направлены на значительное улучшение качества макроэкономических прогнозов и поддержки выработки мер политики в Кыргызской Республике.

Для повышения качества макроэкономических прогнозов, которые в настоящее время разрабатываются Министерством экономики Кыргызской Республики (МЭ КР), в настоящей работе предлагается методика построения модели прогнозирования на *h-шагов* вперед для кыргызского рынка труда, целевыми переменными в которой являются занятость и безработица, измеряемые Национальным статистическим комитетом (НСК)КР на ежеквартальной основе.

В данном отчете, на основе реализации предложенной методики, также представлены прогнозы на 12 кварталов вперед (на период с 1 кв. 2014 г. по 4 кв. 2016 г.) для следующих переменных:

- общая занятость;
- занятость в сельском хозяйстве;
- занятость в промышленности;
- занятость в строительстве;
- занятость в услугах;
- общая безработица.

Путем построений, также возможно получить прогнозы для:

- общих трудовых ресурсов;
- уровня безработицы.

Ядро предложенной модели прогнозирования состоит из двух наборов стохастических уравнений для прогнозирования численности, как занятых, так и безработных. Первый набор относится ко всем лицам, занятым на секторном уровне, определенным, как функция секторного ВВП, если он значим, и других экономических агрегатов. Это обеспечивает согласованность прогноза переменных рынка труда с квартальным ВВП. Второй набор стохастических уравнений сфокусирован на всех безработных, и, в основном, использует информацию, полученную из прошлых наблюдений за тем же рядом. Численность занятых и безработных прогнозируется стохастическими уравнениями; трудовые ресурсы и уровень безработицы определяются путем использования бухгалтерских тождеств.

Модели прогнозирования работают с оригинальными, т.е., не подвергавшимися коррекции, временными рядами. Для исключения эффекта сезонности, полученные прогнозные оценки были скорректированы с учетом сезонных колебаний, в соответствии с процедурой X-13ARIMA-SEATS.

Настоящий отчет организован следующим образом. В Разделе 2 рассмотрена

общая схема методики, которая использовалась для прогнозирования. Она представляет собой инструкцию для всех будущих действий по прогнозированию.

В Разделе 3 изложены результаты прогнозирования на период 2014-2016 гг. В частности, в Разделе 3.1 рассмотрены вопросы доступности данных и временного охвата переменных рынка труда и набора макроэкономических предикторов, которые потенциально полезны для прогнозирования развития рынка труда. В Разделе 3.2 мы вводим и количественно и качественно оцениваем различные конкурирующие сезонные модели для прогнозирования безработицы (параграф 3.2.1), общей занятости (параграф 3.2.2) и занятости по секторам деятельности (параграф 3.3.1). В Разделе 3.3 представлена сезонная коррекция наборов прогнозируемых значений. Все коды, использованные в вычислениях, приведены в Приложении А.

## 2 Методология

В этом разделе рассматривается общая методология, которую предлагается реализовать для прогнозирования переменных рынка труда.

Придерживаясь последних тенденций, отраженных в литературе в области (см., например, Скалин и Тerasвирта, 2002; Элсби, Хобиджин и Сахин, 2008; Болл, Ли и Лунгани, 2013), предлагается опираться на два главных подхода к прогнозированию спроса и предложения рабочей силы. Первый подход основан на свойствах временных рядов интересующих нас переменных, наряду с индикаторами рынка труда на ближайшую перспективу, если они имеются. Вторым основывается на взаимосвязи между занятостью/безработицей и экономическим ростом, как определяется экономической теорией (например, закон Оукена). Независимо от выбранного подхода, в задаче прогнозирования имеются некоторые общие базовые шаги, которые будут проиллюстрированы в следующем разделе.<sup>1</sup>

### 2.1 Шаги прогнозирования

Пользуясь подходом Хиндмана и Атанасопулоса (2014), определяем пять основных шагов.

- **Шаг 1: Определение задачи.** Сначала необходимо хорошо понять переменные, которые должны прогнозироваться, как они определяются и измеряются. Также важно выяснить, каким образом прогнозы будут использоваться в организации, которой требуются прогнозы. Поэтому, прогнозист должен потратить свое время на обсуждение задачи со всеми, кто будет вовлечен в сбор данных, поддержание баз данных и использование прогнозов для будущего планирования.
- **Шаг 2: Сбор информации: статистические данные.** Всегда имеются, по крайней мере, два вида требуемой информации: (а) временной ряд переменных, которые будут прогнозироваться, и (б) потенциальные индикаторы, задаваемые посредством использования экспертных знаний специалистов, занимающихся местными рынками труда, а также наиболее успешной международной практики.
- **Шаг 3: Исследовательский анализ.** Этот шаг включает анализ характеристик ряда, который будет прогнозироваться, и, в конечном счете, преобразование исходного ряда (например, логарифмическое преобразование). Пониманию свойств ряда всегда помогает отображение в виде графика. Некоторые вопросы, на которые следует ответить, включают следующее: Есть ли повторяющиеся характеристики? Есть ли заметный тренд? Важна ли сезонность? Является ли ряд стационарным? Есть ли какие-либо выбросы в данных, для объяснения которых нужен специалист с экспертными знаниями? Насколько сильны взаимосвязи между другими имеющимися потенциальными индикаторами? Этот предварительный анализ будет рассмотрен далее в параграфе 2.2.

---

<sup>1</sup> Содержание последующих параграфов в значительной степени построено на материалах книги Р.Хиндмана и Г.Атанасопулоса "Прогнозирование: Принципы и практика" (R. Hyndman and G. Athanasopoulos "Forecasting: Principles and Practice", 2014; текст имеется в открытом доступе (OTexts): <http://otexts.com/fpp/>.

- **Шаг 4: Выбор модели, оценка и проверка.** Наилучшая для использования модель зависит от наличия исторических данных и существенности взаимосвязей между прогнозируемой переменной и какими-либо казуальными переменными. Обычной практикой является сравнение двух или трех потенциальных конкурирующих моделей прогнозирования. Каждая модель сама по себе является искусственным построением, которое основано на ряде предположений (явных и неявных), и обычно включает один или несколько параметров, подлежащих оценке. Модели, которые мы рассмотрим, включают:

- регрессионные модели;
- модели ARIMA Бокса-Дженкинса;
- динамические регрессионные модели, которые объединяют регрессионные модели с моделями ARIMA, приводящими к регрессиям с ошибками ARIMA.

В процедуре построения модели важную роль играет также проверка модели. Довольно часто проверка модели состоит из оценки точности подгонки модели (как, например, по статистике  $R^2$ ). Вместе с тем, высокие значения  $R^2$  не гарантируют, что модель пригодна для прогнозирования. Графический анализ остатков - эффективный статистический инструмент для проверки (и, в любом случае, полезно увидеть, были ли предположения правильными, или нет). Различные типы графиков остатков из подогнанной модели дают информацию о соответствии различных аспектов модели. Характер остатков может также помочь принять решение, следует ли включить в модель лаги и/или прошлые изменения в некоторых предикторах.

Выбор и проверка моделей для прогнозирования будут рассмотрены далее в параграфе 2.3.

- **Шаг 5: Использование и оценка прогнозной модели.** После того, как модель (или несколько моделей, если у нас есть конкурирующие модели) была выбрана, и ее параметры оценены, модель «потенциально готова» к вычислению прогнозов.

Однако прежде чем модель будет окончательно готова, она должна быть оценена с точки зрения эффективности прогноза «вне выборки». Эффективность модели можно будет оценить должным образом только после того, как станут доступными данные за прогнозируемый период.

Чтобы помочь в оценке точности прогнозов, был разработан целый ряд методов, которые будут рассмотрены в разделе 2.4.

## 2.2 Исследовательский анализ

Временные ряды обычно демонстрируют различные характеристики (компоненты), которые могут быть выделены. Традиционно, эти характеристики связаны с трендами, сезонностью, циклами и нерегулярной (случайной) компонентой.

- **Тренд.** Тренд имеет место, когда наблюдается долгосрочное увеличение или уменьшение данных. Он не обязательно должен быть линейным. В некоторых случаях мы будем упоминать «тренд, изменяющий направление» – если происходит изменение с растущей тенденции к снижающейся тенденции.

- **Сезонность.** Сезонные компоненты существуют, если ряд находится под влиянием сезонных факторов (например, кварталы года, месяцы или дни недели). Сезонность всегда имеет постоянный и известный период.
- **Цикл.** Циклические компоненты существуют, если данные показывают подъемы и падения, которые не имеют постоянного периода. Продолжительность этих колебаний обычно составляет, по крайней мере, 2 года. Часто бывает трудно отделить цикл от тренда, и общая компонента определяется, как тренд-циклическая.

Заметьте, что, если период не изменяется, и связан с тем или иным аспектом календаря, то характеристика является сезонной. В целом, средняя продолжительность циклов больше, чем продолжительность сезонной компоненты, и абсолютная величина циклов обычно более переменчива, чем величина сезонных компонент.

- **Нерегулярная(случайная) компонента.** Нерегулярная компонента временного ряда – это остаток временного ряда после того, как тренд-циклические и сезонные компоненты (включая календарные эффекты) были удалены. Она соответствует высокочастотным колебаниям ряда.

Соответственно этим компонентам, ряд может быть разложен аддитивным или мультипликативным способом. Если мы принимаем аддитивную модель, ряд может быть записан как

$$y_t = T_t + S_t + E_t,$$

где  $y_t$  – данные периода  $t$ ,  $S_t$  – сезонная компонента в период  $t$ ,  $T_t$  – тренд-циклическая компонента в период  $t$  и  $E_t$  – нерегулярная компонента.

Мультипликативная модель может быть записана как:  $y_t = T_t \times s_t \times e_t$ .

Аддитивная модель полезна, когда сезонная вариация относительно постоянна по времени, в то время как мультипликативная модель – когда сезонная вариация увеличивается с течением времени.<sup>2</sup> Для экономических временных рядов обычно применяются мультипликативные модели.

При использовании моделей для временных рядов, важно удостовериться, что ряд является стационарным. Следует отметить, что стационарные или «стационаризованные» ряды относительно легко прогнозировать.

Стационарный временной ряд – это ряд, свойства которого (среднее, вариация, автокорреляция...) не зависят от времени, в течение которого наблюдается ряд. Поэтому, если ряд стационарный, мы в состоянии получить значащую статистику выборки, такую как средние, вариации и корреляции с другими переменными, и сделать прогноза основе этих статистических свойств, которые в будущем будут такими же, какими они были в прошлом.

Временные ряды с трендами, или с сильной сезонностью не стационарны: тренд и сезонность затрагивают значения временного ряда в разные моменты времени.<sup>3</sup>

<sup>2</sup> Основная команда для разложения ряда в R – "decompose". При использовании аддитивной модели - decompose (название ряда, напечатайте = "additive"). При мультипликативном разложении - decompose (название ряда, напечатайте = "multiplicative").

<sup>3</sup> Например, если ряд последовательно возрастает с течением времени, то среднее выборки и вариация будут расти с увеличением размера выборки, и они всегда будут занижать среднее и вариацию в последующие периоды. При этом если среднее и вариация ряда не четко определены, то они не будут коррелировать с другими переменными. По этой причине необходимо быть осторожным при попытках

График стационарного ряда по времени будет примерно горизонтальным, с постоянной вариацией.

Полезным инструментом, позволяющим судить, стационарен ли ряд, является функция автокорреляции. Автокорреляция означает корреляцию временного ряда с его собственными прошлыми и будущими значениями. Точно так же, как корреляция измеряет степень линейной зависимости между двумя переменными, автокорреляция измеряет линейное соотношение между *запаздывающими* значениями временного ряда.<sup>4</sup> Запаздывание является периодом времени, разделяющим упорядоченные данные, и используется для вычисления коэффициентов автокорреляции. Запаздывающая версия временного ряда вычисляется путем смещения шкалы времени назад на определенное число наблюдений. Таким образом, переменная с запаздыванием определяется следующим образом: предположим, что мы имеем временной ряд  $Y: y_t, t = 1, \dots, T$ ; *запаздывающая переменная* (с величиной лага, равной единице) является переменной временного ряда с наблюдениями:

$$z_t = y_{t-1}, t = 2, \dots, T.$$

В общем случае, запаздывающая версия временного ряда, смещенная назад на  $k$  наблюдений, может быть записана как:

$$W_t = Y_{t-k}, \text{ for } t = k + 1, \dots, T.$$

Автокорреляция порядка  $k$  ( $r_k$ ) определяется как корреляция между  $y_t$  и  $y_{t-k}$ :

$$r_k = \frac{\sum_{t=k+1}^T (Y_t - \text{mean}(Y))(Y_{t-k} - \text{mean}(Y))}{\sum_{t=k+1}^T (Y_t - \text{mean}(Y))^2}$$

Функция автокорреляции (выборки) (ACF или коррелограмма) отображает коэффициенты автокорреляции на вертикальной оси, и значения запаздывания на горизонтальной оси, т.е., она отображает  $r_k$  на оси  $y$  как функцию  $k$  ( $k = 1, \dots, K$ ). Здесь представляет максимальную величину запаздывания (эмпирическое правило: никогда не превышайте значение  $K = 1/4 * T$ ). Графическое представление ACF позволяет получить представление о свойствах временного ряда. Например, положительная автокорреляция может рассматриваться, как специфическая форма постоянства, тенденции системы оставаться в том же состоянии от предыдущего наблюдения к следующему.

Если временной ряд *абсолютно случаен*, то для больших значений  $Tr(k) \approx 0$  для всех ненулевых значений  $k$ . Стационарным временным рядам присуща краткосрочная корреляция, характеризующаяся небольшим количеством больших значений  $r(k)$ , которые имеют тенденцию последовательно уменьшаться, и стремятся к нулю при больших значениях  $k$ . Если временной ряд является нестационарным, например, присутствует сильный тренд, то значения  $r(k)$  не будут уменьшаться до нуля, за исключением случаев больших значений  $k$ . Если временной ряд характеризуется сильными сезонными колебаниями, то коррелограмма также покажет большие значения с той же частотой (например, для квартальных данных при  $k = 4, 8, 12 \dots$ ).

Если временной ряд не является стационарным, то во многих случаях его можно преобразовать в стационарный, пользуясь одним из следующих преобразований:

---

экстраполировать регрессионные модели, подогнанные к нестационарным данным.

<sup>4</sup> Ее также иногда называют запаздывающей корреляцией или корреляцией рядов.

- *Взятие разностей.* Взятие разностей может помочь стабилизировать среднее временного ряда путем удаления изменений в уровне временного ряда, и таким образом устранить тренд и сезонность.

*Первой разностью* временного ряда является серия изменений между последовательными наблюдениями в оригинальном ряде:

$$y'_t = y_t - y_{t-1}$$

*Сезонной разностью* временного ряда является серия различий между каким-либо наблюдением и соответствующим наблюдением в предыдущем году:

$$y'_t = y_t - y_{t-m}$$

где  $m$  = количество сезонов.

- Приведение в логарифмическую форму. Такое преобразование, как логарифмирование, может помочь стабилизировать непостоянство временного ряда. Логарифмическая форма удобна, поскольку поддается толкованию: изменения величины логарифма представляют относительные (или процентные) изменения в оригинальном масштабе. Другая полезная особенность логарифмических преобразований заключается в том, что при их применении значения прогнозов всегда остаются положительными в оригинальном масштабе.

Прогнозы для стационаризованного ряда могут быть «детрансформированы», путем *обратного преобразования* любых математических преобразований, которые были применены ранее, с тем, чтобы получить прогнозы для оригинального ряда.

## 2.3 Выбор модели

Прогнозирование эконометрического временного ряда может осуществляться с использованием спектра различных моделей, включающего, в основном, линейные и нелинейные модели. На основе предположения, что временной ряд является реализацией вероятностного процесса, в большинстве прикладных подходов нашел применение линейный класс моделей, поскольку они дают хорошее приближение процесса, производящего данные.

Как показывает практика, к наиболее успешным одномерным моделям для прогнозирования переменных рынка труда относятся:

- (линейные) регрессионные модели (например, модель авторегрессионного распределенного лага);
- модели временных рядов (например, модель ARIMA);
- комбинации этих двух моделей: основываются на моделях временных рядов (модели ARIMA), но включают также переменные предикторов.



### 2.3.1 Линейная регрессионная модель

В линейной регрессионной модели предполагается, что прогнозируемые переменные и переменные предикторов связаны простой линейной моделью:

$$y = \beta_0 + \beta_1 x + \varepsilon,$$

где параметры  $\beta_0$  и  $\beta_1$  определяют, соответственно, точку на оси координат и наклон линии. Точка  $\beta_0$  представляет прогнозируемое значение  $y$  при  $x = 0$ . Наклон  $\beta_1$  представляет прогнозируемое увеличение  $y$  при увеличении  $x$  на одну единицу. Заметьте, что наблюдения не лежат на прямой линии, но разбросаны вокруг нее. Мы можем рассматривать каждое наблюдение  $y_i$ , как состоящее из систематической или объясненной части модели  $\beta_0 + \beta_1 x_i$ , и случайной «ошибки»  $\varepsilon_i$ . «Ошибочный член» подразумевает не ошибку, а отклонение от основной модели прямой линии. Он включает в себя все, что может повлиять на  $y_i$ , кроме  $x_i$ . Предполагается, что эти ошибки:

1. имеют среднее равное нулю; в противном случае прогнозы будут содержать систематическую ошибку;
2. не автокоррелируются; иначе прогнозы будут неэффективны, поскольку в данных содержится больше информации, которая должна быть использована;
3. не связаны с переменной предиктора; иначе в систематическую часть модели было бы необходимо включить больше информации.

На практике у нас есть набор наблюдений, но мы не знаем значений  $\beta_0$  и  $\beta_1$ . Их необходимо оценить на основе данных. Принцип наименьших квадратов обеспечивает способ эффективного выбора  $\beta_0$  и  $\beta_1$  путем минимизации суммы квадратических ошибок. Таким образом, выбираем минимизирующие значения  $\beta_0$  и  $\beta_1$

$$\sum_{i=1}^N \varepsilon_i^2 = \sum_{i=1}^N (y_i - \beta_0 - \beta_1 x_i)^2$$

В результате получаются известные оценочные формулы обычного метода наименьших квадратов (МНК):

$$\hat{\beta}_1 = \frac{\sum_{i=1}^N (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

и

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x},$$

где  $\bar{x}$  - среднее арифметическое  $x$  наблюдений, и  $\bar{y}$  - среднее арифметическое  $y$  наблюдений. Полученная оценочная линия известна, как «линия регрессии».

Прогнозируемые значения  $y$ , полученные из наблюдавшихся значений  $x$ , называются «подобранные значения». Мы записываем их, как  $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ , для  $i = 1, \dots, N$ . Каждое значение  $\hat{y}_i$  представляет собой точку на линии регрессии, соответствующую наблюдению  $x_i$ .

Разница между наблюдавшимися значениями  $y$  и соответствующими подобранными значениями является «остатками»:  $e_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$ .

Общая форма множественной линейной регрессии:

$$y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \dots + \beta_k x_{k,i} + e_i$$

где  $y$  – прогнозируемая переменная,  $x_1, \dots, x_k$  – переменные предиктора  $k$ . Коэффициенты  $\beta_1, \dots, \beta_k$  измеряют влияние каждого предиктора после учета влияния всех остальных предикторов, включенных в модель. Таким образом, коэффициенты измеряют ограниченные воздействия переменных предикторов.

Когда имеется множество возможных предикторов, нам нужна некоторая стратегия выбора наилучших предикторов для включения в модель. Обычный подход заключается в изучении характера корреляции между переменными и возможными предикторами, так как корреляция и регрессия сильно связаны. Преимущество регрессионной модели по сравнению с корреляцией состоит в том, что она предполагает наличие прогнозирующей взаимосвязи между этими двумя переменными ( $x$  прогнозирует, не обязательно являясь причиной  $y$ ), и количественно определяет ее таким способом, который полезен для прогнозирования. Другой общий подход состоит в исключении из полной модели со всеми предикторами тех переменных,  $p$ -значения которых ниже 0,05. Эта стратегия не работает корректно, если два или больше предиктора сильно коррелируются друг с другом, из-за проблемы мультиколлинеарности (которая возникает, когда две или более переменных предикторов дают аналогичную информацию). Эти подходы должны сопровождаться некоторыми мерами обеспечения точности прогнозирования.

После выбора переменных регрессии и подгонки модели регрессии, важным шагом, которого нельзя избежать, является построение графиков остатков, определяемых как разница между фактическими значениями и подобранными значениями  $e_t = y_t - \hat{y}_t$ , чтобы проверить, были ли удовлетворены предположения о модели. Необходимо построить целый ряд графиков, чтобы проверить различные аспекты подогнанной модели и основных предположений. Помимо более общих графиков, таких, как диаграмма рассеяния остатков по времени или по подобранным значениям, если мы имеем дело с данными временного ряда, более важным является график остатков функции автокорреляции (ACF). Этот график показывает, существует ли какая-либо автокорреляция в остатках, что является явным признаком того, что присутствует информация, которая не была учтена в модели. Формальный тест, разработанный для выявления значительной общей автокорреляции, называют тестом портманто (от французского слова, означающего чемодан, который может содержать много разных вещей). Один из таких тестов – тест Льюнга-Бокса, основанный на следующей статистической формуле:

$$Q^* = T(T+2) \sum_{k=1}^h (T-k)^{-1} r_k^2$$

Большие значения  $Q^*$  указывают на то, что причиной автокорреляции не является ряд белого шума.

Желательно всегда проверять, является ли распределение остатков нормальным. Этот шаг не обязателен для прогнозирования, но он значительно облегчает вычисление интервалов прогнозирования.

Всякий раз, когда мы вводим в регрессионную модель данные временного ряда, мы должны принять во внимание так называемую *ложную регрессию* (Грейнджер и Ньюболд, 1974), которая является следствием того, что оценка параметров модели с помощью МНК приводит к вводящим в заблуждение статистически значимым результатам. Рассмотрим, например, стандартную (статическую) регрессию:  $y_t = \beta_0 +$

$\beta_1 x_t + \varepsilon_t$ . Если эти две переменные не связаны, т.е. значение  $\beta_1 = 0$ , то мы должны ожидать что его оценка по МНК  $\hat{\beta}_1 \approx 0$ . Вместо этого МНК может привести к оценке  $\hat{\beta}_1$ , существенно отличной от нуля, и обычные статистические тесты укажут, что коэффициент статистически отличается от нуля.

Эта проблема привела к значительной работе над тем, как определить, какими свойствами должен обладать временной ряд, если будут использоваться эконометрические методы. Основной вывод состоит в том, что любой временной ряд, используемый в эконометрических приложениях, должен быть стационарным. Как правило, временные ряды с экономическими данными не стационарны; то есть, значения временного ряда не колеблются относительно постоянного среднего или с постоянной вариацией (если это следствие стохастического тренда, то они имеют единичный корень).<sup>5</sup> За одним исключением, вы никогда не должны производить регрессию  $Y$  по  $X$ , если переменные имеют единичные корни. Исключением является ситуация, когда  $Y$  и  $X$  являются коинтегрированными. В любом случае, безопасное решение, которое работает во всех случаях, состоит в том, что сначала берутся разности ряда, а затем осуществляется вычисление по регрессионной модели.

### 2.3.2 Модели ARIMA

Модели ARIMA используют только информацию о переменной, которая будет прогнозироваться, и не делают попыток определить факторы, которые влияют на ее поведение. В связи с этим, они могут экстраполировать тренд и сезонные компоненты, однако игнорируют всю другую информацию, такую как изменения в экономических условиях, которые могут быть отражены предикторами в регрессионной модели.

Прежде чем подробно рассматривать модели ARIMA, следует ввести некоторые обозначения.

Оператор обратного смещения  $B$ — удобный механизм<sup>6</sup> для работы с запаздыванием во временных рядах:

$$By_t = y_{t-1}.$$

Другими словами,  $B$ , воздействующий на  $y_t$ , создает эффект смещения данных назад на один период. Применение  $B$  к  $y_t$  дважды смещает данные назад два периода:

$$B(By_t) = B^2 y_t = y_{t-2}$$

Для квартальных данных, если мы хотим рассмотреть «тот же квартал прошлого года», используется оператор

$$B_4 y_t = y_{t-4}$$

Оператор обратного смещения удобен для описания процесса взятия разностей.

Первая разность может быть записана как:

<sup>5</sup> Существует ряд тестов на единичные корни, которые основаны на разных предположениях и могут приводить к противоречивым ответам. Одним из наиболее популярных является Улучшенный тест Дики-Фуллера (ADF).

<sup>6</sup> В некоторых источниках используется  $L$  для обозначения “лага” вместо  $B$  для “обратного смещения”.

$$y'_t = y_t - y_{t-1} = y_t - By_t = (1 - B)y_t$$

Обратное смещение также очень полезно при объединении разностей, поскольку с оператором можно действовать, используя обычные алгебраические правила. В частности, условия, содержащие  $B$ , могут перемножаться. Например, сезонную разность, на которую накладывается первая разность, можно записать, как:

$$(1 - B)(1 - B^m)y_t = (1 - B - B^m + B^{m+1})y_t = y_t - y_{t-1} - y_{t-m} + y_{t-m-1}.$$

## Авторегрессионные (АР) модели

В авторегрессионной модели мы прогнозируем интересующую нас переменную, используя линейную комбинацию прошлых значений переменной. Термин авторегрессия указывает на то, что это - регрессия переменной по самой себе.

Таким образом авторегрессионная модель порядка  $p$  может быть записана как:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t,$$

где  $c$  - константа, и  $a_t$  белый шум. Это напоминает множественную регрессию, но с запаздывающими значениями  $y_t$ , в качестве предикторов. Данная модель называется АР ( $p$ ).

Авторегрессионные модели имеют необыкновенную гибкость при обработке широкого диапазона различных характеристик временных рядов. Изменение параметров  $\phi_1 \dots \phi_p$  приводит к различным характеристикам временного ряда. Вариация члена ошибки  $a_t$  изменяет только масштаб ряда, но не характеристики.

Для модели АР(1):

- если  $\phi_1 = 0$ ,  $y_t$  эквивалентен белому шуму;
- если  $\phi_1 = 1$  и  $c = 0$ ,  $y_t$  эквивалентен случайному блужданию;
- если  $\phi_1 = 1$  и  $c \neq 0$ ,  $y_t$  эквивалентен случайному блужданию со смещением;
- если  $\phi_1 < 0$ ,  $y_t$  имеет тенденцию колебаться между положительными и отрицательными значениями.

Ограничиваем применение авторегрессионных моделей стационарными данными, что приводит к некоторым ограничениям на значения параметров:

- для модели АР(1):  $-1 < \phi_1 < 1$ .
- для модели АР (2):  $-1 < \phi_2 < 1$ ,  $\phi_1 + \phi_2 < 1$ ,  $\phi_2 - \phi_1 < 1$ .
- если  $p \geq 3$  ограничения становятся намного более сложными.

## Модели со скользящим средним (СС)

Вместо использования прошлых значений прогнозируемой переменной в регрессии, модель скользящего среднего (СС) использует прошлые ошибки прогноза в модели, подобной регрессии:

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q},$$

где  $e_t$  - белый шум. Мы называем эту модель **СС(q)**. Конечно, мы не наблюдаем значения  $e_t$ , поэтому она не является регрессией в обычном смысле. В этих моделях каждое значение  $y_t$  может считаться взвешенным скользящим средним значением небольшого числа прошлых ошибок прогноза. Изменение параметров  $\theta_1, \dots, \theta_q$  приводит к различным характеристикам временного ряда. Как и в случае авторегрессионных моделей, вариация ошибки  $e_t$  изменяет только масштаб ряда, но не характеристики.

Заметьте, что это можно написать любую стационарную модель AP (p) как модель СС ( $\infty$ ).

Например, используя повторные замены, мы можем продемонстрировать это для модели AP (1):

$$\begin{aligned} y_t &= \phi_1 y_{t-1} + e_t \\ &= \phi_1(\phi_1 y_{t-2} + e_{t-1}) + e_t \\ &= \phi_1^2 y_{t-2} + \phi_1 e_{t-1} + e_t \\ &= \phi_1^3 y_{t-3} + \phi_1^2 e_{t-2} + \phi_1 e_{t-1} + e_t \end{aligned}$$

И т.д.

При условии, что  $-1 < \phi_1 < 1$ , значение  $\phi_1^k$  будет уменьшаться по мере увеличения  $k$ . Таким образом, в конечном счете, мы получим:

$$y_t = e_t + \phi_1 e_{t-1} + \phi_1^2 e_{t-2} + \phi_1^3 e_{t-3} + \dots,$$

процесс СС( $\infty$ ).

Инвертируемые модели обладают некоторыми математическими свойствами, которые облегчают их использование в практике. Ограничения инвертируемости аналогичны ограничениям стационарности:

- для модели СС(1):  $-1 < \theta_1 < 1$ ;
- для модели СС(2):  $-1 < \theta_2 < 1$ ,  $\theta_1 + \theta_2 < 1$ ,  $\phi_2 - \theta_1 < 1$ ;
- для  $q \geq 3$  используются более сложные условия.

С точки зрения оценки, модель СС невозможно оценить обычным МНК – требуются другие методы.

### Модели ARIMA без учета сезонной компоненты

Если мы объединим разности с авторегрессией и моделью скользящего среднего, мы получим несезонную модель ARIMA. ARIMA - акроним названия модели Auto Regressive Integrated Moving Average (авторегрессионная интегрированная со скользящим средним; «интеграция» в этом контексте означает действие, обратное взятию разностей). Полная модель может быть записана как:

$$y'_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} + e_t, \quad (1)$$

где  $y'_t$  - разностный ряд (разности могут последовательно браться несколько раз).

«Предикторы» в правой части включают, как значения  $y_t$  с запаздыванием, так и ошибки с запаздыванием.

Данная модель называется ARIMA (p, d, q), где:

- $p$  = порядок авторегрессионной части;
- $d$  = порядок взятия разностей;
- $q$  = порядок части скользящего среднего значения.

Выбор соответствующих значений для  $p$ ,  $d$  и  $q$  не является легкой задачей. Общепринятая практика – выполнение процедуры Бокса-Дженкинса, представленной на рисунке 1. Обратите внимание на то, что ACF и функция частной автокорреляции (PACF) полезны при выборе порядка модели ARIMA.<sup>7</sup> Автоматизированный алгоритм выбора порядка модели ARIMA имеется в R: функция `auto.arima()` в прогнозной библиотеке.

Важно отметить, что константа  $c$  оказывает важное воздействие на долгосрочные прогнозы, полученные с использованием этих моделей:

Если  $c=0$  и  $d=0$ , долгосрочные прогнозы будут сходиться к нулю.

Если  $c=0$  и  $d=1$ , долгосрочные прогнозы будут сходиться к константе, отличной от нуля.

Если  $c=0$  и  $d=2$ , долгосрочные прогнозы будут представлять прямую линию.

Если  $c \neq 0$  и  $d=0$ , долгосрочные прогнозы будут сходиться к среднему значению данных.

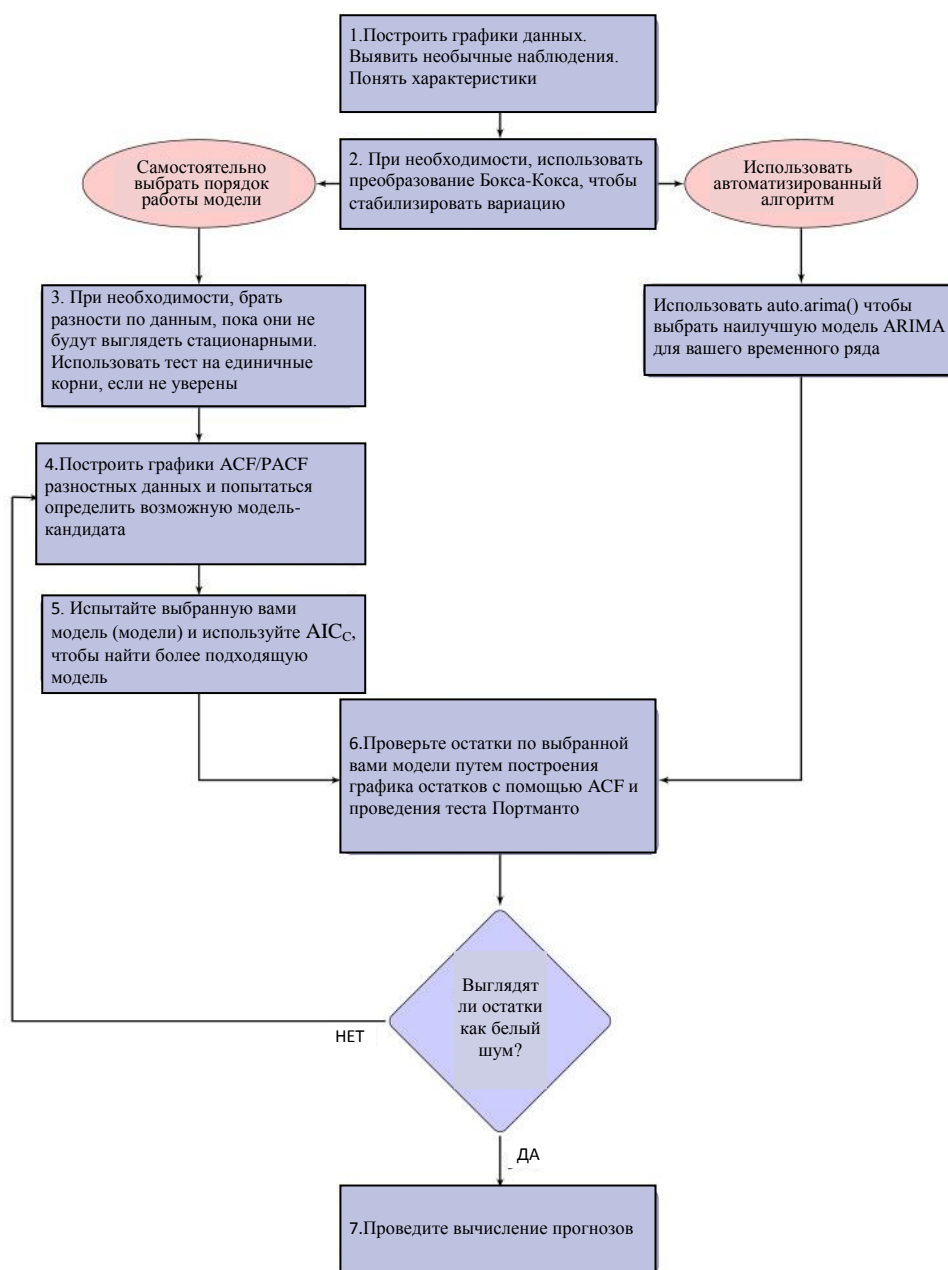
Если  $c \neq 0$  и  $d=1$ , долгосрочные прогнозы будут представлять прямую линию.

Если  $c \neq 0$  и  $d=2$ , долгосрочные прогнозы будут повторять квадратический тренд.

---

<sup>7</sup> Вспомните, что график ACF показывает автокорреляции, которые измеряют линейную зависимость между  $y_t$  и  $y_{t-k}$  для различных значений  $k$ . Если  $y_t$  и  $y_{t-1}$  коррелируются, то  $y_{t-1}$  и  $y_{t-2}$  должны также коррелироваться. Но тогда  $y_t$  и  $y_{t-2}$  могут коррелироваться, просто потому что они оба связаны с  $y_{t-1}$ , а не из-за какой-то новой информации, содержащейся в  $y_{t-2}$ , которая могла бы использоваться в прогнозировании  $y_t$ . Чтобы преодолеть эту проблему, мы можем использовать частичные автокорреляции. Они измеряют зависимость между  $y_t$  и  $y_{t-k}$  после удаления эффектов других временных лагов: 1, 2...,  $k-1$ .

**Рисунок 1:** Общий процесс прогнозирования с использованием модели ARIMA. Рисунок из книги Хиндмана и Атанасопулоса (2014 г.).



## Сезонные модели ARIMA

Сезонная модель ARIMA формулируется путем введения дополнительных сезонных условий в модели ARIMA, с которыми мы уже познакомились.

Сезонная часть модели AP или CC проявляется в виде сезонных лагов функций автокорреляции. При выборе соответствующих порядков для сезонной части модели ARIMA ограничьте внимание сезонными лагами. Процедура моделирования почти не отличается от процедуры для несезонных данных, за исключением того, что мы должны выбрать сезонную AP и условия CC, а также несезонные компоненты модели.

Это записывается следующим образом:

Рисунок 2: ARIMA

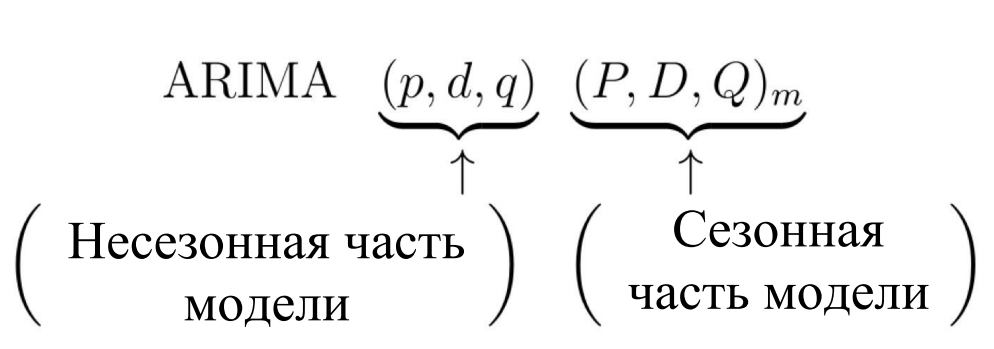


Рисунок 3: Сезонная  $ARIMA(1,1,1) \times (1,1,1)$

$$(1 - \phi_1 B) (1 - \Phi_1 B^4) (1 - B) (1 - B^4) y_t = (1 + \theta_1 B) (1 + \Theta_1 B^4) e_t.$$

The equation is annotated with arrows pointing to each factor, indicating its component type:

- $(1 - \phi_1 B)$  is labeled "Несезонная AP(1)" (Non-seasonal AP(1)).
- $(1 - \Phi_1 B^4)$  is labeled "Сезонная AP(1)" (Seasonal AP(1)).
- $(1 - B)$  is labeled "Несезонная разница" (Non-seasonal difference).
- $(1 - B^4)$  is labeled "Сезонная разница" (Seasonal difference).
- $(1 + \theta_1 B)$  is labeled "Несезонная CC(1)" (Non-seasonal CC(1)).
- $(1 + \Theta_1 B^4)$  is labeled "Сезонная CC(1)" (Seasonal CC(1)).



### 2.3.3 Регрессионные модели с ошибками ARIMA

Регрессионные модели допускают включение полезной информации из переменных предикторов, но не допускают сложную динамику временных рядов, которая может обрабатываться в моделях ARIMA. Другой класс моделей, которые могут использоваться в прогнозировании, является общим классом динамических регрессионных моделей. Начнем с простого объединения регрессионных моделей и моделей ARIMA, чтобы получить регрессию с ошибками ARIMA.

Начнем с регрессионных моделей формы:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \dots + \beta_k x_{k,t} + e_t$$

где прогнозируемая переменная  $y$  является линейной функцией переменных  $x_1, \dots, x_k$  предиктора  $k$ , и  $e_t$  – член, представляющий ошибки, который считается некоррелирующимся (например, белый шум).

Здесь, однако, мы разрешим, чтобы ошибки из регрессии содержали автокорреляцию.

Чтобы подчеркнуть это изменение в перспективе, заменим в уравнении  $e_t$  на  $n_t$ .

Предположим, что ряд ошибок  $n_t$  соответствует модели ARIMA. Например, если  $n_t$  соответствует модели ARIMA(1,1,1), мы можем записать:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \dots + \beta_k x_{k,t} + n_t,$$
$$(1 - \phi_1 B)(1 - B)n_t = (1 + \theta_1 B)e_t,$$

где  $e_t$  – ряд белого шума.

Кроме того, при оценке регрессии с ошибками ARMA, в первую очередь необходимо, чтобы все переменные в модели были стационарными. Таким образом, мы сначала должны убедиться, что  $y_t$  и все предикторы, судя по всему, являются стационарными. Если мы проводим оценку на модели, но при этом какие-либо из них нестационарны, оцениваемые коэффициенты могут быть неправильными.

Если все переменные в модели стационарны, то в качестве остатков нам остается рассматривать только ошибки ARMA. Регрессионная модель с ошибками ARIMA эквивалентна регрессионной модели в разностях, с ошибками ARMA. Например, если взять разности от вышеупомянутой регрессионной модели с ошибками ARIMA (1,1,1), мы получаем модель

$$y'_t = \beta_1 x'_{1,t} + \beta_2 x'_{2,t} + \dots + \beta_k x'_{k,t} + n'_t$$
$$(1 - \phi_1 B)n'_t = (1 + \theta_1 B)e_t$$

где

$$y'_t = y_t - y_{t-1}; x'_{i,t} = x_{i,t} - x_{i,t-1}; n'_t = n_t - n_{t-1}$$

что является регрессионной моделью в разностях с ошибками ARMA.

Функция `Аrima()` в R вводит в регрессионную модель ошибки ARIMA, если используется аргумент `xreg`. Аргумент порядка определяет порядок модели с ошибками ARIMA. Если задано взятие разностей, то разности берутся по всем переменным в регрессионной модели до проведения оценки с помощью модели.

В некоторых случаях, воздействие предиктора, включенного в регрессионную

модель, может иметь временное смещение. В таких ситуациях мы должны допускать возможность лага воздействия предиктора. Предположим, что у нас есть только один предиктор в модели. Тогда модель, которая допускает воздействия с лагом, может быть записана, как

$$y_t = \beta_0 + \gamma_0 x_t + \gamma_1 x_{t-1} + \dots + \gamma_k x_{t-k} + n_t$$

где  $n_t$  - процесс ARIMA.

### 2.3.4 Эмпирическая стратегия: краткий обзор

Вкратце, эмпирическая стратегия выбора модели для прогнозирования переменных рынка труда включает следующее:

- Анализ свойств временного ряда прогнозируемых переменных.
- Определение порядка интеграции, которая может привести к необходимости преобразования требуемых оригинальных данных, прежде, чем проводить оценки на модели (первая разность, сезонная разность, логарифмическое преобразование...).
- Определение предикторов (с любым лагом) на основе корреляций и исследовательской регрессии.
- Если ни один из предикторов не имеет значительной связи с прогнозируемой переменной → реализовать модель ARIMA, используя процедуру Бокса-Дженкинса.
- Если есть значительные предикторы (с любым лагом) → реализовать регрессионную модель.
- Если диагностика остатков регрессионной модели дала хорошие результаты → использовать регрессионную модель для прогнозирования.
- Если множественная регрессия показывает плохие результаты диагностики (например, остатки по-прежнему автокоррелируются), реализовать регрессионную модель с ошибками ARMA.
- Стараться сохранять модели максимально простыми!

## 2.4 Оценка прогнозирования

Задача прогнозирования состоит в предсказании будущего с максимальной точностью, на основе всей имеющейся информации, включая исторические данные и знание любых будущих событий, которые могут повлиять на прогнозы.

Наша цель состоит в том, чтобы прогнозировать будущие значения экономических переменных (таких, как занятость), поэтому от нас требуется оценка модели прогнозирования, не обязательно с привлечением экономической теории, но с максимальным вниманием нахождению реалистических значений.

В этом отношении важно не перепутать корреляцию с причинной обусловленностью, или причинную обусловленность с прогнозированием. Переменная  $x$  может быть полезна для прогнозирования переменной  $y$ , но это не означает, что  $x$  является причиной  $y$ . Даже если  $x$  является причиной  $y$ , взаимосвязи между ними могут быть более сложными, чем простая причинная связь. Корреляции полезны для прогнозирования, даже если причинная связь между этими двумя переменными отсутствует.

Цифры, которые мы пытаемся предсказать, неизвестны (иначе мы бы их не прогнозировали!), и таким образом, мы можем рассматривать их, как **случайные переменные**. Например, общая занятость в следующем квартале может находиться в некотором диапазоне возможных значений, и пока мы не просуммируем фактическую занятость в конце квартала, мы не знаем, каково будет ее значение. Поэтому, пока мы не знаем занятость в следующем квартале, она остается случайной величиной.

Получая прогноз, мы оцениваем середину диапазона возможных значений, которые могла принять случайная переменная. Очень часто прогноз сопровождается интервалом предсказания, дающим диапазон значений, которые случайная переменная может принимать с относительно высокой вероятностью. Например, 95%-й интервал предсказания содержит диапазон значений, которые должны включать фактическое будущее значение с вероятностью 95%.

Прогноз всегда основан на той информации, которая имеется на момент прогнозирования. Предположим, что мы обозначили всю информацию, которую мы наблюдали во время  $t$  как  $INF_t$ , и хотим предсказать  $y_{t+1}$ . Мы можем записать  $y_{t+1}/INF_t$ , что означает «случайную переменную  $y_{t+1}$ , с учетом того, что мы знаем  $INF_t$ ». Набор значений, которые эта случайная переменная могла принять, наряду с их относительными вероятностями, известен как «распределение вероятности»  $y_{t+1}/INF_t$ . В прогнозировании его называют распределением прогноза.

Когда мы говорим о «прогнозе», мы обычно имеем в виду среднее значение распределения прогноза, и помещаем «крышку» над  $y$ , чтобы показать это:  $\hat{y}_{t+1}|t$ . Когда мы прогнозируем временной ряд, мы можем сделать другую запись:  $F_{t+h} = F_{t,h}$ , указывающую, что это прогноз на период  $t + h$ , основанный на информации, доступной в момент времени  $t$ ;  $h$  - так называемый, горизонт прогноза, и  $F_{t,h}$  - прогноз  $h$ -шагов вперед.

В большинстве ситуаций прогнозирования, по мере приближения события, вариация величины, которую мы предсказываем, уменьшается. Другими словами, чем дальше вперед мы предсказываем, тем большую неопределенность мы имеем.

Обратите внимание на то, что модель, которая хорошо соответствует данным, не обязательно дает хороший прогноз. Важно оценить точность прогноза, используя подлинные данные. Не достаточно убедиться в том, что модель хорошо соответствует историческим данным. Точность прогнозов может быть определена только путем рассмотрения, насколько хорошо модель показывает себя с новыми данными, которые не использовались при подгонке модели.

## Тренировочная совокупность

При оценке эффективности прогнозирования, обычно используют часть доступных данных для подгонки (*тренировочные тесты*, сформированные из  $T - s$  наблюдений, например, для  $t$  от 1 до  $T - s$ ), а остальная часть данных используется для тестирования модели (*тестовые наборы*, сформированные из  $s$  наблюдений, например, для  $t$  от  $T - s + 1$  до  $T$ ). Поэтому, тестовые данные могут использоваться для оценки того, насколько хорошо модель будет прогнозировать при использовании новых данных. Размер тестового набора, как правило, составляет приблизительно 20% всей выборки, хотя это значение зависит от того, насколько продолжительна выборка, и насколько далеко вперед Вы хотите предсказывать. Размер тестового набора в идеальном случае должен быть, по крайней мере, настолько большим, насколько велик требуемый максимальный горизонт прогноза. *Тренировочный набор* также называют «данными в выборке», а *тестовый набор* - «данными вне выборки».

Для оценки прогноза используется несколько мер. Их можно разделить на

масштабно-зависимые и масштабнo-независимые меры. Пусть  $y_i$  обозначает  $i$ -тое наблюдение, а  $\hat{y}_i$  обозначает значение прогноза, равное  $y_i$ . Ошибка прогноза выражается просто как  $e_i = y_i - \hat{y}_i$ , и находится в том же масштабе, что и данные.<sup>8</sup>

### Масштабно-зависимые

Меры точности, которые основаны на  $e_i$ , зависимы от масштаба и не могут использоваться для сравнений между рядами, которые измеряются различных масштабах.

Две наиболее распространенные масштабнo-зависимые меры основаны на абсолютных ошибках или квадратических ошибках:

Средняя абсолютная ошибка:  $MAE = \text{mean}(|e_i|)$ ,

Среднеквадратическая ошибка:  $RMSE = \sqrt{\text{mean}(e_i^2)}$

### Масштабно-независимые

Наиболее распространенной масштабнo-независимой мерой является средняя абсолютная процентная ошибка (MAPE), которая основана на ошибках в процентах. Процентная ошибка определяется как  $p_i = 100e_i/y_i$ . Преимуществом процентных ошибок является их независимость от масштаба, поэтому они часто используются при сравнении эффективности прогнозов между различными наборами данных.

MAPE определяется как:

средняя абсолютная процентная ошибка:  $MAPE = \text{mean}(|p_i|)$

Недостаток мер, основанных на процентных ошибках, заключается в том, что они имеют бесконечное или неопределенное значение, если  $y_i = 0$  для любого  $i$  в интересующий нас период, и экстремальные значения, когда какое-либо  $y_i$  близко к нулю.

Средняя абсолютная масштабированная ошибка (MASE), предложенная Хиндманом и Кехлером (2006), представляет хорошую альтернативу использованию ошибок в процентах при сравнении точности прогнозов между рядами в различных масштабах. Они предложили масштабировать ошибки на основе тренировочных MAE, полученной из простого метода прогнозирования. Удобным способом определения масштабированной ошибки является использование наивных прогнозов<sup>9</sup>:

$$q_j = \frac{e_j}{e^{naive}}$$

Где для несезонных временных рядов (простейший случай),  $e^{naive}$  определяется как:

$$e^{naive} = \frac{1}{s} \sum_{j=1}^s |y_{T-s+j} - y_{T-s}|$$

а для сезонных временных рядов, как:

<sup>8</sup> Во временных рядах мы обозначаем  $y_t$  меру переменной  $y$  в момент времени  $t$ .

<sup>9</sup> Наивный прогноз просто получить, задав все прогнозы равными значению последнего наблюдения. Т.е., прогнозы всех будущих значений задаются равными  $y_T$ , где  $y_T$  - последняя наблюдавшаяся величина

$$\overline{e^{naive}} = \frac{1}{s} \sum_{j=1}^s |y_{T-s+j} - y_{T-s}^m|$$

где  $y_{T-s}^m$  – последнее наблюдавшееся значение в тренировочном наборе для того же сезона  $y_{T-s+j}$ . Поскольку и числитель и знаменатель содержат значения в масштабе оригинальных данных,  $q_j$  независим от масштаба данных. Масштабированная ошибка будет меньше единицы, если она является результатом лучшего прогноза, чем средний наивный прогноз, вычисленный по тренировочным данным. С другой стороны, она будет больше единицы, если прогноз хуже, чем средний наивный прогноз, вычисленный по тренировочным данным.

Средняя абсолютная масштабированная ошибка определяется как

$$MASE = \text{mean}(|q_j|)$$

Аналогично, средняя квадратическая масштабированная ошибка (MSSE) может быть определена путем возведения в квадрат ошибок (из тренировочных и тестовых данных), вместо использования их абсолютных значений.

### 3 Прогнозирование результатов для периода 2014–2016 гг.

#### 3.1 Обработка данных

Прогноз будет основан на имеющихся данных, используя не только ежеквартальные ряды по рынку труда, предоставляемые Национальным статистическим комитетом (НСК) в Обследовании занятости и безработицы, но также и макроэкономические индикаторы, представляемые в Национальных счетах. В целом, наша основная стратегия состояла в том, чтобы максимизировать использование всей доступной информации, в попытке связать как можно больше индикаторов рынка труда с макропеременными Национальных счетов.

Первый набор данных состоит из рядов занятости и безработицы, полученных из Обследования занятости и безработицы<sup>10</sup>, проводимого Отделом статистики труда и занятости Национального статистического комитета<sup>11</sup> Кыргызской Республики. Он также включает квартальные ряды среднемесячной зарплаты и безработных, зарегистрированных государственными органами занятости.

В частности, наша разработка была выполнена на основе квартальных данных по занятости и безработице, начиная с первого квартала 2004 г., и до последних доступных данных на момент реализации модели, т.е., последнего квартала 2013 г. Более подробно – в наш анализ были включены, как общая занятость, так и ее разбивка по главным секторам экономики: имеющиеся данные позволили произвести разделение по четырем главным секторам экономической деятельности, т.е., сельскому хозяйству, промышленности, строительству и услугам. Модели были построены для необработанных данных, т.е. без коррекции на сезонные колебания.

Определение занятости и безработицы соответствовало определению, данному в интегрированном выборочном обследовании бюджетов домашних хозяйств и рабочей силы.

Вторая часть информации состоит из некоторых макроиндикаторов, полученных из Национальных счетов Кыргызской Республики (НС КР)<sup>12</sup>. В частности после тщательного рассмотрения всего набора предоставленных индикаторов, мы составили выборку потенциальных ключевых переменных, которые было бы полезно интегрировать в структуру прогнозирования. В результате был составлен окончательный набор макропеременных, которые включают:

- (i) совокупный ВВП(объем);
- (ii) добавленная стоимость по основным экономическим секторам(объем);
- (iii) частное потребление(объем);
- (iv) общественное потребление (объем);
- (v) валовые инвестиции (объем);
- (vi) экспорт товаров и услуг(объем);
- (vii) импорт товаров и услуг(объем).

---

<sup>10</sup> Первый выпуск модуля «Занятость и безработица» Национального статистического комитета (НСК) состоялся в ноябре 2002 г. в рамках проекта “ Укрепления статистического потенциала в поддержку процесса ККР/ПРСР”, финансируемого Министерством международного развития Великобритании (DFID) и совместно реализованного НСК и "Oxford Policy Management" (OPM).

<sup>11</sup><http://www.stat.kg>

<sup>12</sup> Национальные счета составляются НСК Кыргызской Республики в соответствии с основными концепциями, категориями, определениями и другими методологическими принципами Системы национальных счетов - 1993. Данные составляются на ежеквартальной основе и выпускаются через 90 дней после окончания квартала. Дополнительную информацию см.: <http://dsbb.imf.org/pages/sdds/DQAFBase.aspx?ctycode=KGZ&catcode=NAG00>.

Для получения агрегатов в реальном выражении, т.е. в объемах, номинальные агрегаты были скорректированы по ценам путем использования соответствующих дефляторов из НСК КР. Данные о дефляторах публикуются ежеквартально, и представляют собой индексы цен агрегатов относительно предыдущего периода (например, соответствующего квартала предыдущего года). Агрегаты в «постоянных ценах» были получены методом «линковки цепочек», который включает следующие шаги:<sup>13</sup>

- вычисляются агрегаты в ценах предыдущего периода (агрегат в текущих ценах, деленный на соответствующий дефлятор);
- вычисляются изменения в агрегатах между предыдущим и текущим периодом в ценах предыдущего периода (в результате этой процедуры получаем индекс объемов Ласпейреса для периода между  $t - 4$  и  $t$ );
- эти изменения соединяются в цепочку (т.е., каждое умножается на последующее) отдельно по каждому агрегату (что в результате дает цепочечный индекс объема);
- и, наконец, чтобы обеспечить ряды объема на уровнях, ряд умножается на значение агрегата в текущих ценах соответствующего квартала базового года, которым, в нашем случае, является 2004 год.

Обратите внимание на то, что эта процедура эквивалентна делению агрегата в текущих ценах на цепочный индекс цен (т.е. умножающий каждый дефлятор на последующий). Еще раз отметим, данные были предоставлены на квартальной основе, и не подвергались коррекции для учета сезонности. На момент прогноза они охватывали период с первого квартала 2004 г. до четвертого квартала 2014 г., т.е., на четыре квартала вперед относительно ряда рынка труда.

Исходя из имеющихся данных, можно получить меру производительности труда, определяемую, как мера объема продукции, произведенной на одного наемного работника. Мера объема продукции здесь определяется, как валовая добавленная стоимость в постоянных ценах.

---

<sup>13</sup> Леквильер Ф. и Блэйдс Д. (2014), Понимание национальных счетов, 2-е издание, ОЭСР.

## 3.2 Квартальные прогнозы рынка труда на период 2014-2016 г.

### 3.2.1 Безработица

Поквартальный ряд показателя безработицы охватывает период с 2004 по 2013 год. Мы хотим построить прогнозы безработицы на 12 кварталов вперед. Необработанные квартальные данные за период 2004–2013 гг. показаны на рисунке 4. Приближенное разложение ряда безработицы на тренд-цикл, сезонность и случайную компоненту показано на рисунке 5.

Общая безработица в Кыргызстане с 2004 по 2013 г.—Квартальные данные

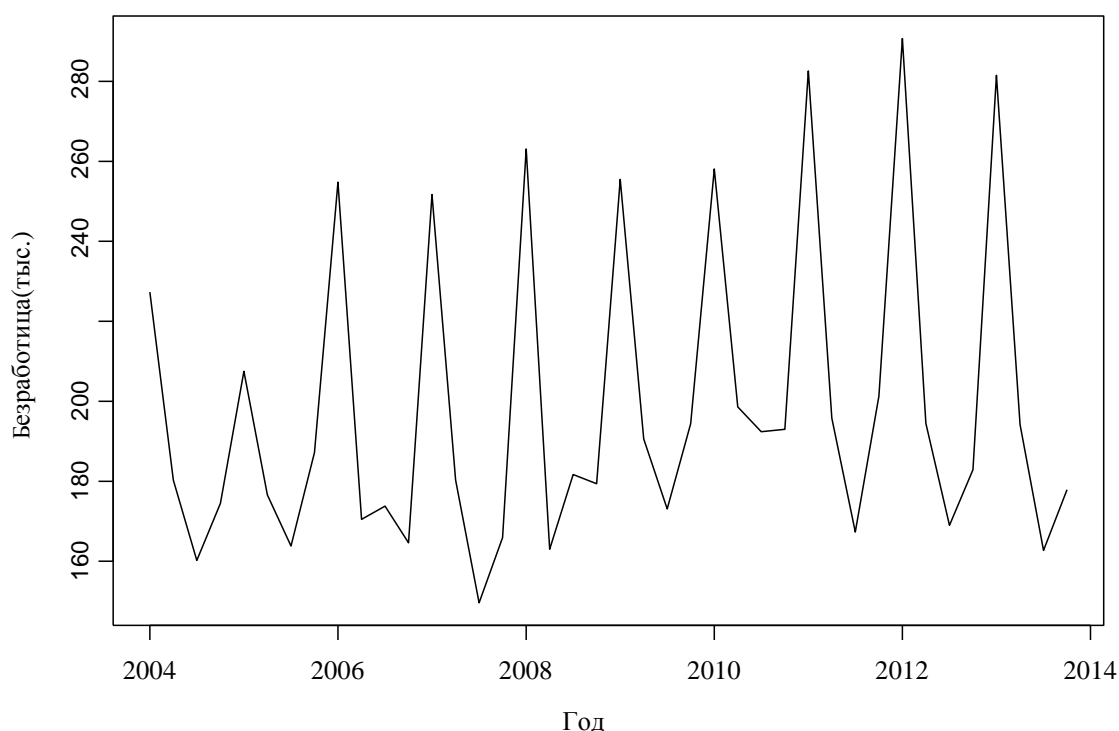


Рисунок 4: Квартальный уровень безработицы в Кыргызстане в период с 2004 по 2013 г.

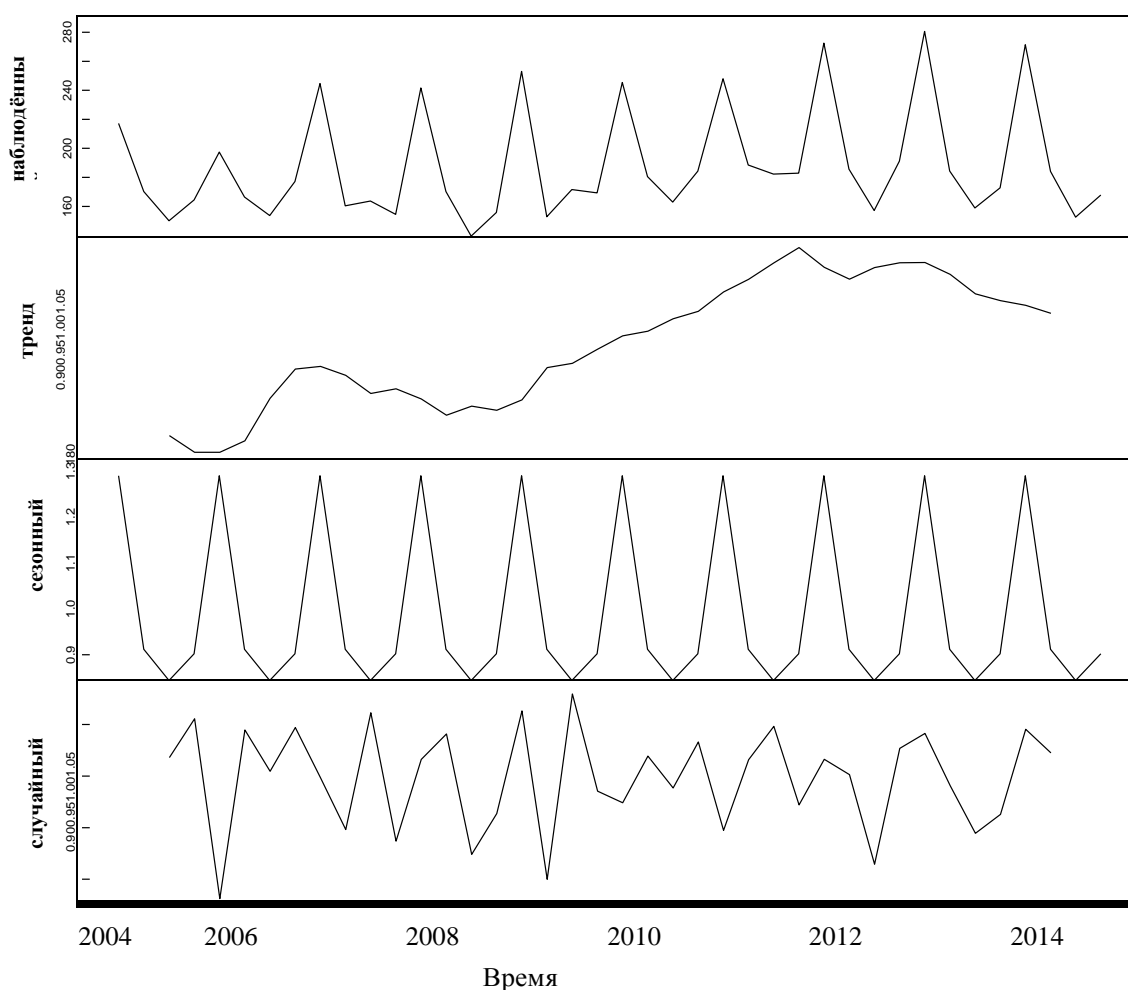
Данные показывают сильную постоянную сезонность, что подтверждается автокорреляционными функциями (см. рисунок 6). Поэтому, начинаем вычитать сезонную разность. В нашем случае сезонная разность - разница между наблюдением и соответствующим наблюдением за такой же квартал предыдущего года:

$$y_t' = y_t - y_{t-4}.$$

Поскольку мы вычитаем наблюдение, сделанное с лагом в 4 периода, такую разность также называют «разностью с лагом -4». Данные со взятой сезонной разностью показаны на рисунке 7, вместе с функциями автокорреляции.



### Разложение мультипликативного временного ряда



**Рисунок5:** Мультипликативное разложение ряда безработицы на ее компоненты

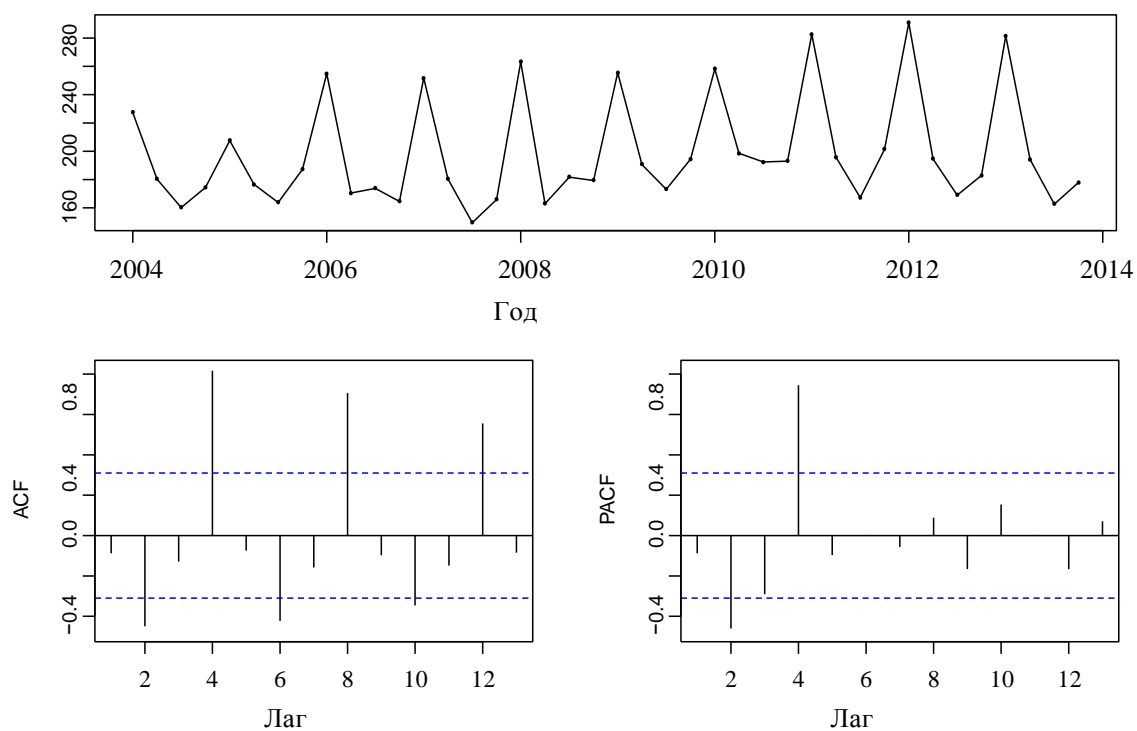
Данные, судя по всему, достаточно стационарны, и поэтому не нуждаются в дополнительном раунде взятия разностей. Однако всегда существует определенная степень субъективной оценки в выборе разностей, которые следует применить. Хотя это и не имеет особого значения, в тех случаях, когда ряд имеет сильный сезонный характер, рекомендуется сначала взять сезонную разность, потому что получившийся ряд может быть стационарным, и, таким образом, отпадет необходимость в дополнительных процедурах взятия разностей.

Поскольку отсутствует определенность в отношении стационарности ряда, мы переходим к взятию разностей второго порядка:

$$y_t'' = y_t' - y_{t-1}' = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}.$$

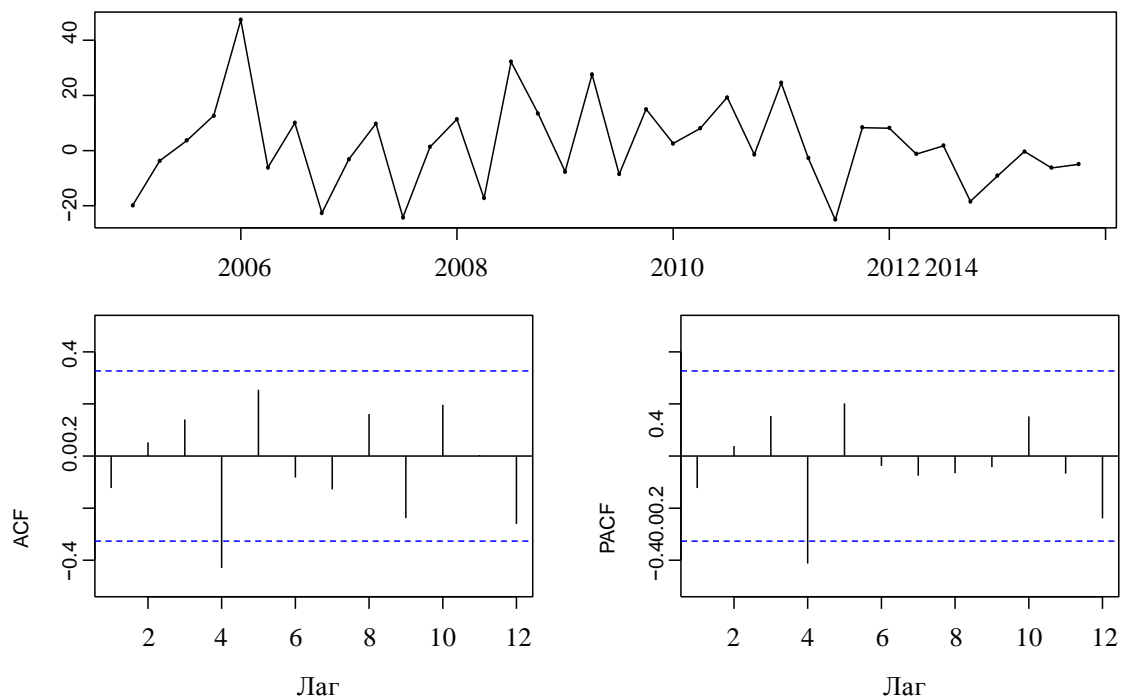
Данные после взятия разностей второго порядка показаны на рисунке 8.

### Ряд безработицы и его автокорреляционные функции

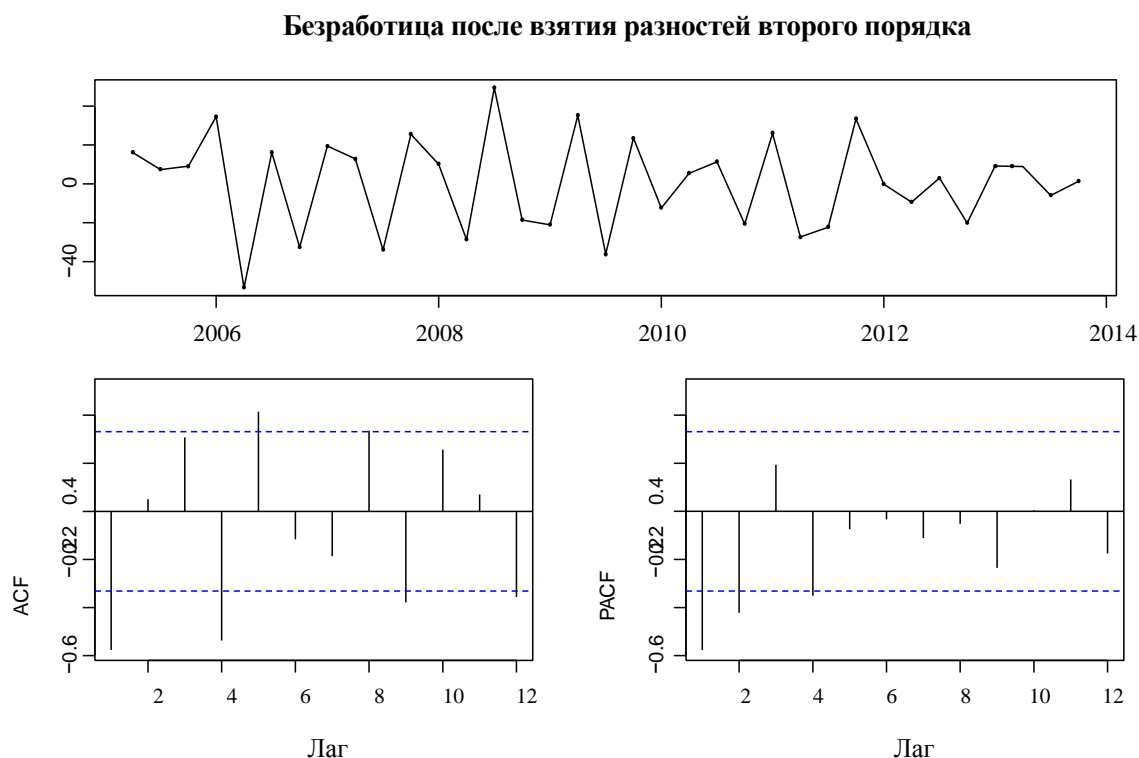


**Рисунок 6:** *Квартальные уровни бедности в Кыргызстане в период с 2004 по 2013 г. и функции автокорреляции*

### Безработица после взятия сезонной разности



**Рисунок 7:** *Безработица в Кыргызстане в период с 2004 г. по 2013г. после взятия сезонной разности*



**Рисунок 8:** Безработица в Кыргызстане в период с 2004 по 2013 г. после взятия разностей второго порядка.

Из верхней диаграммы на рисунке 8 видно, что ряд является стационарным, однако существует риск излишнего взятия разностей, которое приведет ко многим статистически значимым автокорреляциям. Это становится очевидным при рассмотрении соответствующих автокорреляционных функций на нижних диаграммах рисунка 8.

Оба порядка разностей в нашем случае могут быть приняты во внимание при прогнозировании в демонстрационных целях.<sup>14</sup>

Структура кросс-корреляции для квартальных изменений ряда показывает (см Таблицу 1), что имеющиеся макроиндикаторы плохо коррелируют в разностях с рядом безработицы с любым лагом. Поэтому, они не вносят полезного вклада в объяснение временного характера изменений безработицы. Дальнейший анализ с регрессом сезонных изменений безработицы по сезонным изменениям макроэкономических переменных не показал присутствия значимых коэффициентов. Поэтому, прогнозирование безработицы базируется исключительно на внутренних свойствах самого ряда. Таким образом, модель для этой переменной будет сезонной моделью ARIMA.

Выбор сезонной модели ARIMA учитывает оба варианта взятия разностей:

$$\begin{aligned} d = 0 \text{ и } D = 1; \\ d = 1 \text{ и } D = 1. \end{aligned}$$

Значительный пик в лаге 4 автокорреляционных функций на рисунке 7 указывает на сезонный компонент CC (1), который должен быть включен в сезонную

<sup>14</sup> Для того чтобы более объективно определить, требуется ли взятие разностей, можно провести формальные тесты. Тем не менее, всегда остается некоторая степень неопределенности, и некоторые решения должны приниматься аналитиком.

часть модели. Таким образом, мы начинаем подгонять следующие модели: ARIMA (0, 0, 0) (0, 1, 1)<sub>[4]</sub>, и ARIMA (0, 1, 0) (0, 1, 1)<sub>[4]</sub>, наряду с одной вариацией ARIMA (0, 1, 1) (0, 1, 1)<sub>[4]</sub>. В конечном счете, мы добавили еще модель с дрейфом ARIMA (0, 0, 0) (0, 1, 1)<sub>[4]</sub>, которая включает составляющую дрейфа/тренда, с тем, чтобы увидеть, будет ли среднее ряда смещаться со временем немного вверх, как это можно предположить из изучения исходных данных. Фактически, даже при том, что взятие разности удаляет тренды, у модели ARIMA (0, 0, 0) (0, 1, 1)<sub>[4]</sub>, может быть тренд, если она содержит константу, которая отлична от нуля и равна среднему данных после первого взятия разности.

**Таблица1:** Структура кросс-корреляции квартальных изменений между безработицей и потенциальными макроэкономическими предикторами.

Лаг	Зан. пром	ВВП	ДС С/х	ДС пром	ДС Стр	ДС Усл	Част. потреб	Об. потреб	Инвест
-4	-0.429	0	0.049	-0.057	0.117	-0.039	-0.006	-0.088	0.048
-3	0.14	0.012	0.011	-0.313	-0.138	-0.113	0.003	0.113	0.03
-2	0.053	-0.144	-0.169	0.427	-0.085	-0.232	-0.268	0.081	-0.126
-1	-0.123	-0.015	0.05	-0.131	-0.193	-0.099	0.179	0.008	0.003
0	1	-0.288	-0.212	0.029	-0.279	-0.161	-0.271	0.083	0.095
Лаг	Занят	Эксп	Имп	Зарпл	Произ. всего	Произ. с/х	Произ. прм	Произ. з.стр	Произ. з.усл
-4	0.429	0.002	-0.073	0.171	-0.069	-0.074	0.069	0.03	0.035
-3	0.14	0.136	0.186	0.012	-0.128	-0.369	0.011	-0.215	-0.068
-2	0.053	0.032	-0.097	-0.225	-0.035	0.553	-0.183	-0.142	-0.272
-1	-0.123	0.06	0.192	-0.025	-0.115	0.008	-0.003	-0.296	-0.196
0	1	-0.233	-0.226	-0.039	-0.166	0.263	-0.23	-0.285	-0.261

После проведения оценки<sup>15</sup>, все эти четыре модели проходят диагностику остатков и автокорреляционный тест Льюнга-Бокса, направленный на подтверждение случайности остатков. Поэтому, все эти модели - возможные кандидаты на прогнозирование.

«Лучшая» модель из числа имеющихся будет выбрана на основе скорректированного информационного критерия Акаике (AIC<sub>c</sub>)<sup>16</sup> и их эффективности прогнозирования, что включает оценку спектра сводных мер точности прогноза, как, например, среднеквадратическая ошибка между наблюдаемыми данными и прогнозными значениями в тестовом наборе (вне выборки).

Таким образом, мы проводим тест на потенциальную эффективность

<sup>15</sup> Обратите внимание, что в R, при оценке модели ARIMA, используется оценка максимального правдоподобия (ОМП). Эта методика находит значения параметров, которые максимизируют вероятность получения данных, которые мы наблюдали. Для моделей ARIMA, ОМП очень близка к оценкам методом наименьших квадратов. Отметьте также, что модели ARIMA намного более сложны для оценки, чем регрессионные модели, и применение различных видов программного обеспечения даст несколько отличающиеся ответы, поскольку они используют различные методы оценки или различные алгоритмы оценки.

<sup>16</sup> Информационный критерий Акаике - мера относительного качества статистических моделей для данного набора данных. При наличии коллекции моделей для данных, AIC оценивает качество каждой модели относительно каждой из остальных моделей. Таким образом, AIC используется, как средство для выбора модели. Его можно записать, как  $AIC = -2\log(L) + 2(p + q + k + 1)$ , где  $L$  - правдоподобие данных, а  $k=1$  указывает на присутствие дрейфа. Обратите внимание на то, что последний элемент в круглых скобках - число параметров в модели (включая вариацию остатков).

прогнозирования имеющихся конкурирующих моделей. С этой целью имеющийся ряд разделяется на набор тренировочных данных для оценки модели, и набор тестовых данных для оценки прогноза. Размер набора тестовых данных, как правило, составляет приблизительно 20% всей выборки, даже если это значение зависит от протяженности горизонта прогноза, и оно не должно быть больше, чем тренировочная часть ряда.

Обратите внимание на то, что конкурирующие модели, которые будут сравниваться с использованием значений  $AIC_c$ , должны иметь одинаковый порядок взятия разностей. Это требование подразумевает, что из четырех имеющихся моделей, только для трех моделей может проводиться сравнение значений  $AIC_c$  (см. Таблицу 2).

С точки зрения подгонки, модель с  $d = 0$  и  $D = 1$  с дрейфом, судя по всему, является лучшей, по сравнению с другими моделями-кандидатами тем же порядком взятых разностей. Вычисленные значения  $AIC_c$  показаны в Таблице 2. Модель  $ARIMA(0, 0, 0)(0, 1, 1)_{[4]}$  с дрейфом также проходит все остаточные тесты (со значением  $p$  в тесте Льюнга-Бокса, равным 0,76).

**Таблица 2:** Сравнение конкурирующих моделей с сезонной компонентой  $CS$  с одинаковым порядком взятых разностей: скорректированный информационный критерий Акаике.

Модель	$AIC_c$
$ARIMA(0,0,0)(0,1,1)_{[4]}$	300.87
$ARIMA(1,0,1)(0,1,1)_{[4]}$	304.01
$ARIMA(0, 0, 0)(0, 1, 1)_{[4]}$ с дрейфом	267.35

Модель, которая хорошо соответствует данным, не обязательно обеспечивает хороший прогноз. Поэтому, мы продолжим наше упражнение, и сравним наши оригинальные четыре конкурирующие модели, используя тестовый набор.<sup>17</sup> Первоначально используем в качестве тестового набора последние восемь кварталов ряда. Таким образом, мы подгоним все выбранные модели, используя данные с 1-го кв. 2004 по 4-й кв. 2011 года, и получим прогнозы количества безработных на период с 1-го кв. 2012 по 4-й кв. 2013 года. Результаты для мер точности прогнозирования представлены в Таблице 3.

**Таблица 3:** Оценка с помощью тестового набора: сравнение проводится по среднеквадратической ошибке (RMSE), средней абсолютной масштабированной ошибке (MASE) и средней абсолютной ошибке в процентах (MAPE).

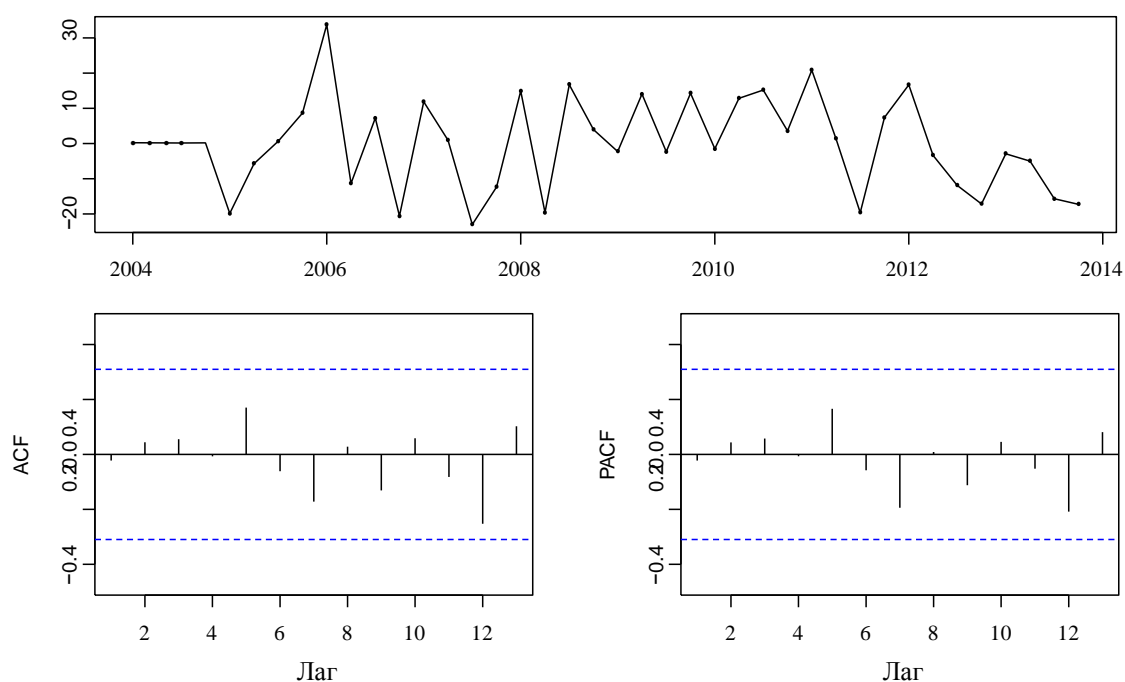
Модель	RMSE	MASE	MAPE
$ARIMA(0,0,0)(0,1,1)_{[4]}$	11.900	0.695	4.769
$ARIMA(1,0,1)(0,1,1)_{[4]}$	18.187	1.165	8.589
$ARIMA(0,1,1)(0,1,1)_{[4]}$	18.920	1.209	8.982
$ARIMA(0, 0, 0)(0, 1, 1)_{[4]}$ с дрейфом	19.796	1.284	9.458

Результаты в Таблице 3 приводят нас к выбору модели  $ARIMA(0, 0, 0)(0, 1, 1)_{[4]}$

<sup>17</sup> Для сравнения моделей с использованием тестового набора данных не обязательно, чтобы модели имели одинаковый порядок взятия разностей. Поэтому, в данном случае мы можем также включить модель с  $d = 1$ .

без дрейфа, которая также проходит все остаточные тесты (с р-значением в тесте Льюнга-Бокса, равным 0,70). Два различных критерия выбора модели не приводят к какой-либо одной модели (сравните результаты Таблицы 2 и Таблицы 3), и тогда требуются дальнейшие исследования. Если мы имеем дело с кратковременным рядом, то тренировочный и тестовый наборы не достаточно длинны, чтобы сделать стабильные выводы. Одно возможное решение (Хиндман и Атанасопулос, 2013) состоит в том, чтобы использовать «перекрестную проверку временного ряда», которая, по существу, заключается в использовании многих различных тренировочных наборов, каждый из которых содержит на одно наблюдение больше, чем предыдущий, при этом соответствующий тестовый набор сокращается на одно наблюдение. Меры точности прогнозирования вычисляются по каждому тестовому набору, и результаты сравниваются. Мы выполняем это упражнение для моделей ARIMA (0, 0, 0)(0, 1, 1)[4] с дрейфом и без дрейфа, являющихся последними конкурентами. Меры точности прогноза, полученные при каждом моделировании, приводят к выбору модели ARIMA (0, 0, 0)(0, 1, 1)[4] с дрейфом. Ее остатки представлены на рисунке 9. Все пики теперь находятся внутри пределов значимости, и таким образом, остатки, судя по всему, являются белым шумом. Тест Льюнга-Бокса также показал, что остатки не имеют оставшихся автокорреляций.

**Остатки подогнанной модели ARIMA (0,0,0)(0,1,1)[4] с дрейфом**



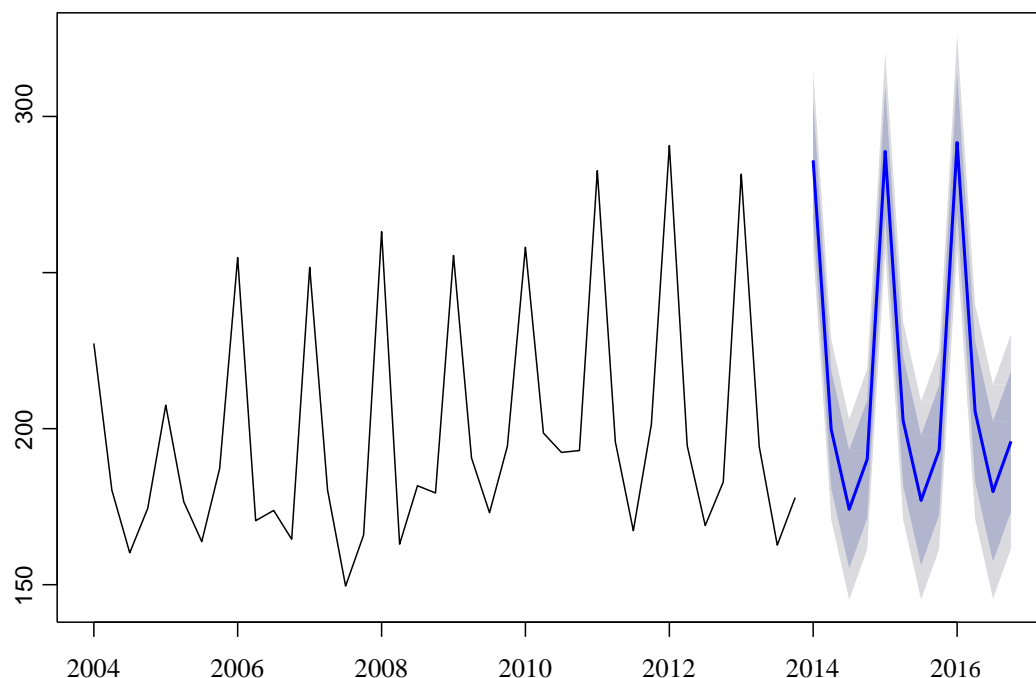
**Рисунок 9:** Остатки подогнанной модели ARIMA (0,0,0)(0,1,1)<sub>[4]</sub> с дрейфом.

Наконец, оцененная сезонная модель ARIMA:

$$\Delta_4 y_t = y_t - y_{t-4} = 0.712 + a_t - 0.548a_{t-4}$$

готова для прогнозирования. Прогнозы, построенные с помощью модели, на последующие три года, показаны на рисунке 10.

**Прогнозы безработицы с использованием модели (0,0,0)(0,1,1) [4] с дрейфом**



**Рисунок 10:** Прогнозы уровня безработицы с использованием модели  $ARIMA(0, 0, 0)(0, 1, 1)_{[4]}$  с дрейфом, с интервалами предсказаний 80% и 95%.

Прогнозные значения, а также их 80% и 95% интервалы предсказаний, представлены в Таблице 4.

**Таблица 4:** Прогнозные значения уровня безработицы (тыс.) с использованием модели  $ARIMA(0, 0, 0)(0, 1, 1)_{[4]}$  с дрейфом, с интервалами предсказаний 80% и 95%.

Квартал	Прогноз	Ниж.80	Верх.80	Ниж.95	Верх.95
2014 кв.1	285.9	267.0	304.9	257.0	314.9
2014 кв.2	199.8	180.9	218.7	170.9	228.7
2014 кв.3	174.1	155.2	193.1	145.2	203.1
2014 кв.4	190.2	171.3	209.1	161.3	219.2
2015 кв.1	288.8	268.0	309.5	257.1	320.5
2015 кв.2	202.6	181.9	223.4	170.9	234.4
2015 кв.3	177.0	156.2	197.8	145.3	208.7
2015 кв.4	193.1	172.3	213.8	161.3	224.8
2016 кв.1	291.6	269.2	314.1	257.3	326.0
2016 кв.2	205.5	183.1	227.9	171.2	239.8
2016 кв.3	179.8	157.4	202.3	145.5	214.2

### 3.2.2 Общая занятость

Квартальный ряд занятости, охватывающий период с 2004 по 2013 г., представлен на рисунке 11. Грубое разложение ряда занятости на тренд-цикл, сезонность и случайный компонент показано на рисунке 12. Ряд, судя по всему, является нестационарным, с сильной сезонностью, как ясно видно из рисунка 13.

Общая занятость в Кыргызстане с 2004 по 2013 г. – квартальные данные

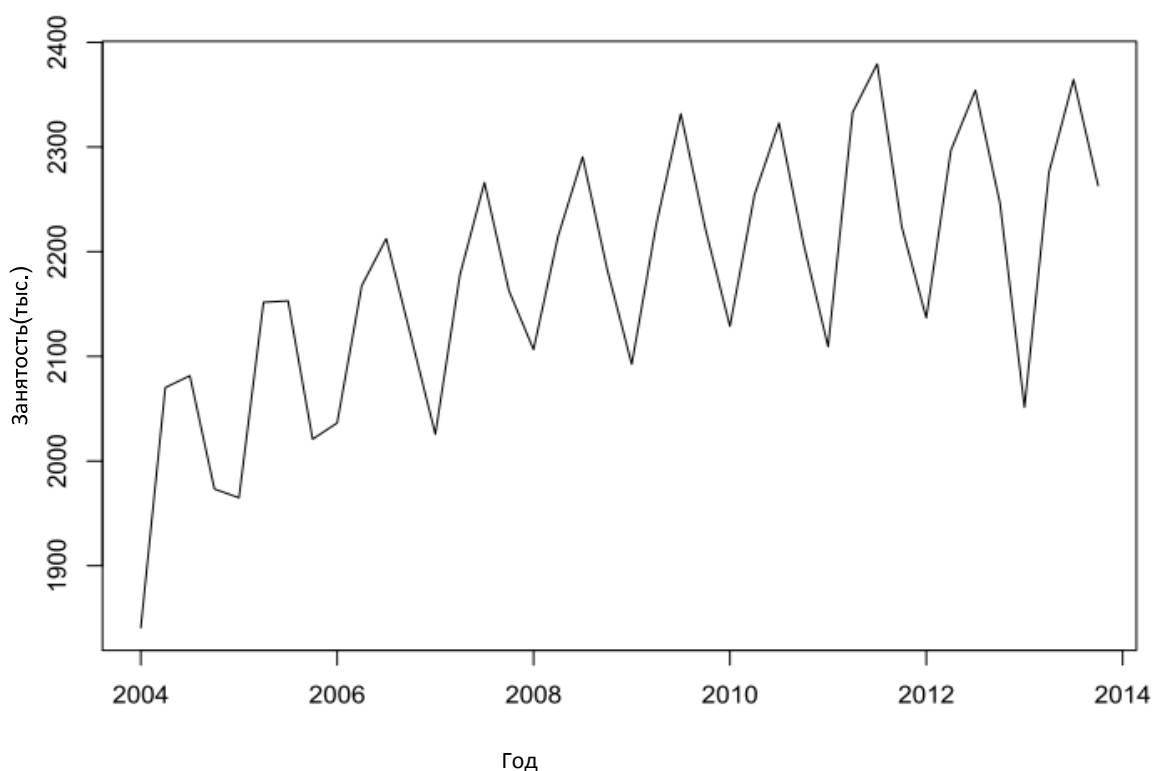


Рисунок 11: Квартальные данные по занятости в Кыргызстане с 2004 по 2013 г.

Ввиду присутствия сезонности, мы, в первую очередь, берем сезонную разность. Данные после взятия сезонной разности показаны на рисунке 14. Теперь ряд выглядит стационарным и таким образом, мы не берем дополнительную первую разность, чтобы избежать чрезмерного взятия разностей.

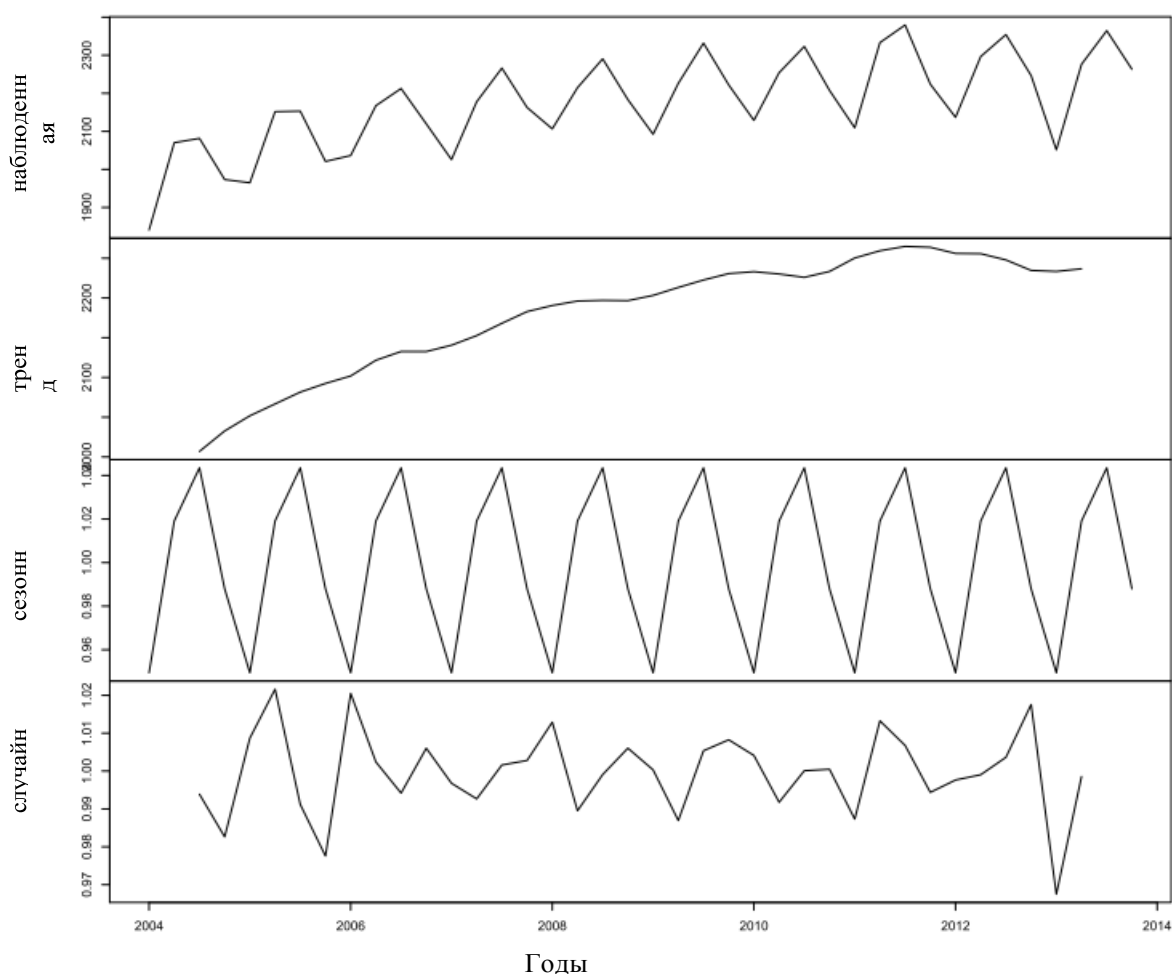
Наши модели-кандидаты для прогнозирования занятости представляют собой модели, в которых зависимость между занятостью, объемом производства, заработной платой и другими макроэкономическими переменными, такими, как производительность, иллюстрируется посредством множества линейных регрессионных моделей, где занятость представлена в виде зависимой переменной, являющейся условной функцией занятости.

Чтобы построить модель для прогнозирования занятости, начнем с рассмотрения зависимости между рядом занятости после взятия разности и макроэкономическими переменными после взятия разности, имеющими экономическое значение в отношении ряда занятости. Обратите внимание на то, что некоторые экономические агрегаты имеются с опережением(ВВП, добавленная стоимость, зарплата..), в то время как другие имеются только на тот же момент времени, что и занятость. В связи с этим, при прогнозировании занятости, необходимо прогнозировать все эти переменные. Требуется более детальное рассмотрение переменных производительности.



Фактически, занятость, произведенный продукт и производительность строго взаимосвязаны.

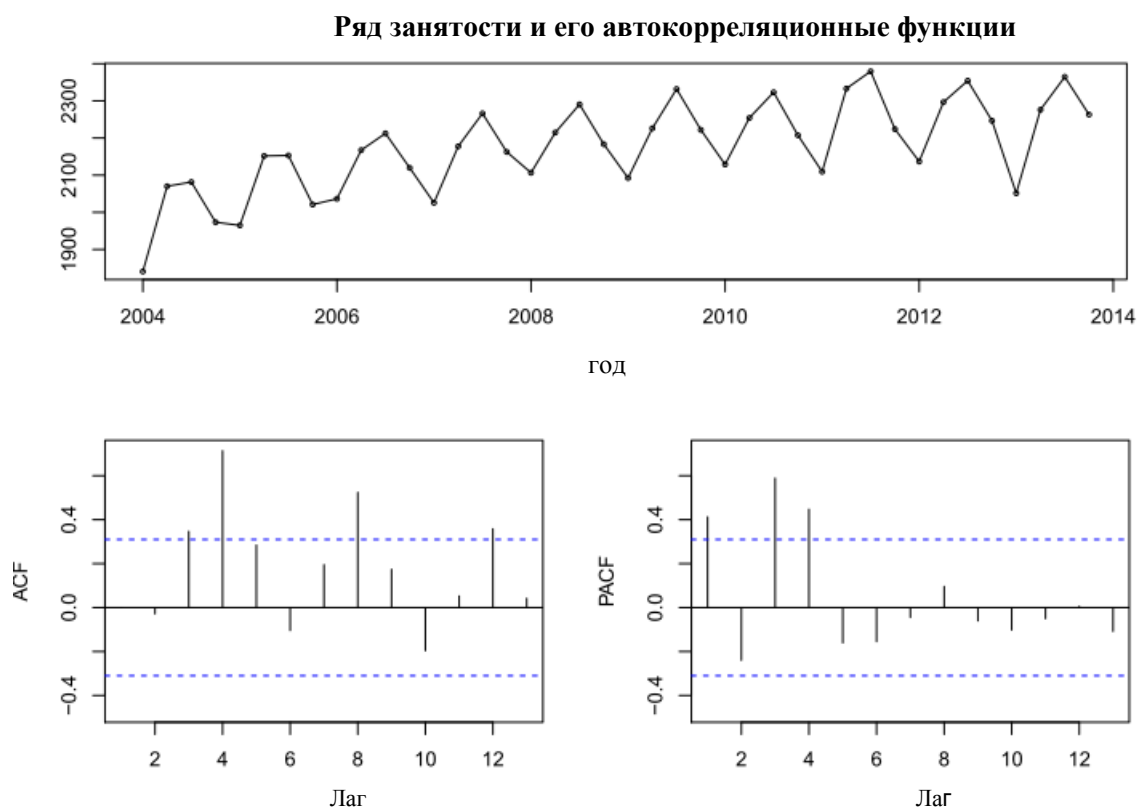
### Разложение мультипликативного временного ряда



**Рисунок 12:** Мультипликативное разложение ряда занятости на компоненты.

Если бы мы могли точно предсказать объем производства и производительность труда, то занятость можно было бы получить просто как соотношение между объемом производства и производительностью. Поэтому, если бы мы использовали переменные производительности и ВВП в качестве предикторов в регрессионной модели занятости, то, вполне возможно, обеспечили бы очень хорошую подгонку. Однако с точки зрения прогнозирования, проблема получения хорошего прогноза потребности в рабочей силе переходит в проблему получения хорошего прогноза производительности. Если для прогнозирования производительности требуется модель ARIMA, то более эффективно применить модель ARIMA непосредственно к ряду занятости. Исходя из этого, настоятельно рекомендуется не использовать производительность как предиктор занятости.

Если ряды нестационарны, рекомендуется производить оценку модели в разностях, для чего необходимо взять разности всех переменных, которые включены в модель, с тем, чтобы сохранить форму зависимости между результатом и предикторами. В Таблице 5 приведены структуры кросс-корреляции между рядом занятости после взятия сезонной разности и рядами квартальных изменений потенциальных предикторов.



**Рисунок 13:** Квартальная занятость в Кыргызстане за период с 2004 по 2013 г. и ее автокорреляционные функции.

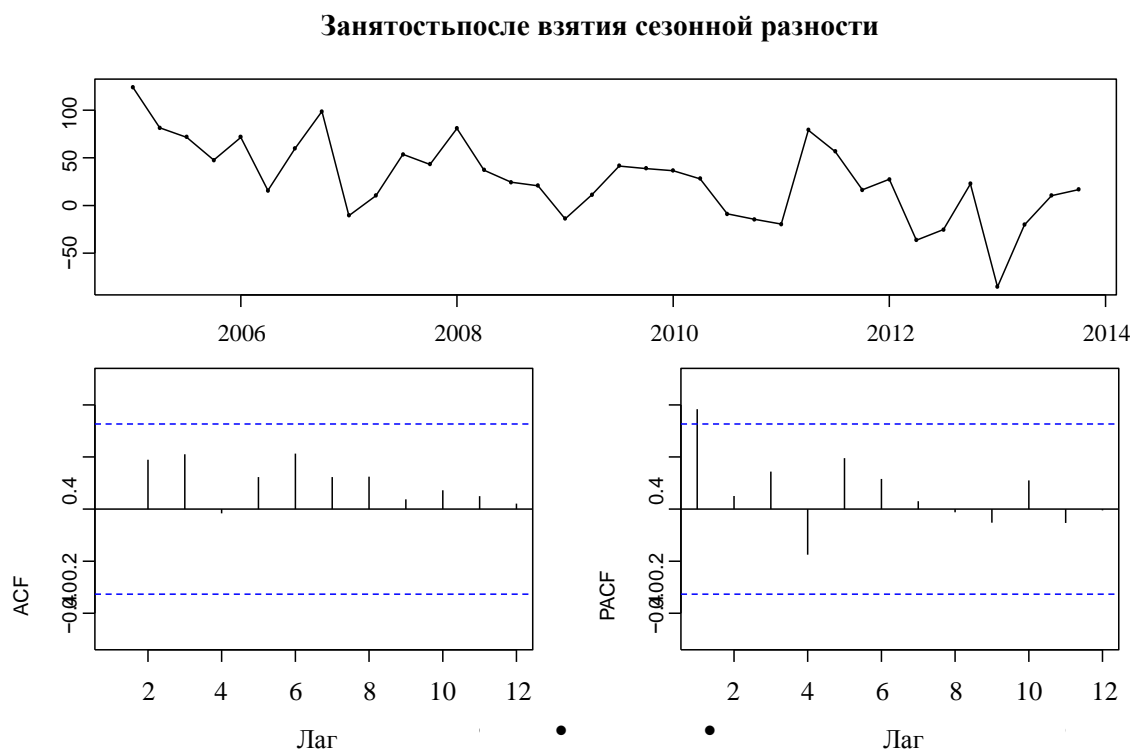
Изучение структуры корреляции ряд после взятия сезонной разности позволяет предположить, что –за исключением некоторых значений занятости с лагами–ВВП, инвестиции, зарплата и производительность, взятые с различными лагами, значительно связаны с занятостью. Обратите внимание на то, что при данном количестве имеющихся наблюдений, мы можем утверждать, что коэффициенты корреляции существенно отличаются от нуля при 5-процентном уровне значимости для значений, больше 0,29 (в абсолютном выражении), и для значений больше 0,4 при 1-процентном уровне значимости. Стратегия, которую мы использовали для выбора наилучших предикторов – это *обратная пошаговая регрессия*. Мы начали с модели, содержащей все потенциальные предикторы, то есть, предикторы после взятия разностей и с различными лагами.

Затем мы по одному исключали предикторы, если они оказывались статистически незначимыми, и если не улучшали точность прогнозирования. По причине, изложенной выше, в модели прогнозирования переменные производительности не принимались во внимание.

В результате обратной пошаговой процедуры, окончательная модель включает, в качестве экзогенных переменных, текущий ВВП после взятия сезонной разности, и инвестиции после взятия сезонной разности с лагом в один период. Диагностика остатков подогнанной модели частично не удовлетворительна, так как остатки все еще автокоррелируются. Затем мы вводим и оцениваем динамическую регрессионную модель для прогнозирования будущей динамики уровней занятости в Кыргызской Республике, которая дает возможность обработки сложной динамики временного ряда с ошибками ARIMA. После того, как мы получаем стационарный ряд, мы должны указать структуру ARMA для ошибок в «модели в

разностях». Фактически, мы теперь разрешаем, чтобы члены, представляющие ошибки из регрессии, содержали автокорреляцию, и, таким образом, придерживаемся модели ARMA.

Как автокорреляционная функция (ACF), так и частная автокорреляционная функция (PACF) ряда после взятия сезонной разности (см. рисунок 14), показывают значительный пик в лаге 1, указывая на наличие некоторых дополнительных компонент в несезонной части модели.



**Рисунок 14:** Занятость после взятия сезонной разности и ее автокорреляционные функции

Таким образом, возможными моделями-кандидатами являются динамические регрессионные модели с ВВП (с лагом 0) и инвестициями (с лагом 1) в качестве внешних предикторов, включающие авторегрессионные ошибки AR(1), либо AR(1) с ошибками дрейфа или ошибки AR(2). Получающиеся значения остатков во всех моделях-кандидатах проходят тест на независимость. Однако дальнейшее изучение показывает, что модель, содержащая члены ошибок AR(1) с дрейфом, имеет минимальное значение AIC<sub>c</sub>, по сравнению с моделями с членами ошибок AR(1) и AR(2) без дрейфа. Кроме того, значимость экзогенной переменной выражена более явно. Исходя из этих соображений, мы продолжаем работать с моделью AR(1) с дрейфом для ошибок. Получающиеся остаточные значения представлены на рисунке 15. Тест Льюнга-Бокса подтверждает, что остатки не коррелируются ( $p$ -значение=0.534).

**Таблица 5:** Структура кросс-корреляции квартальных изменений между занятостью и потенциальными макроэкономическими предикторами.

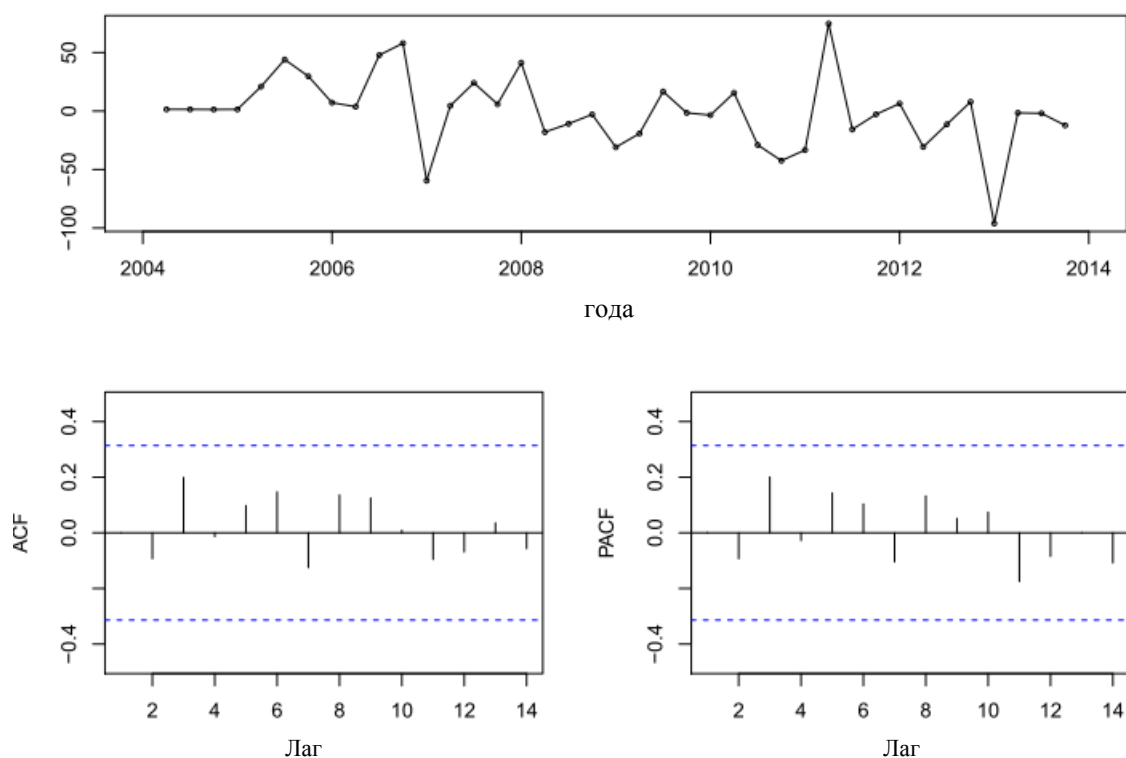
Лаг	Занят	ВВП	ДС. Прм	ДС с/х	ДС стр.	ДС. Усл	Ч. потреб	Об.потреб.	Инв
-4	-0.016	-0.183	-0.326	0.071	0.164	0.205	-0.043	0.035	0.175
-3	0.210	-0.348	-0.392	-0.145	0.066	0.307	0.206	-0.008	0.246
-2	0.189	-0.316	-0.386	-0.037	-0.033	0.295	0.050	0.029	0.254
-1	0.383	-0.044	-0.090	-0.018	0.226	0.209	0.054	0.132	0.282
0	1.000	0.074	0.055	-0.121	-0.025	0.264	0.168	0.007	0.040
Лаг	Занят	Эксп	Имп	Зарпл	Произ. всего	Произ. с/х	Произ.5я ции	Произ. Стр.	Произ. Усл.
-4	-0.016	-0.013	0.140	0.130	-0.109	0.315	-0.364	0.016	0.038
-3	0.210	-0.101	0.315	0.441	-0.336	0.081	-0.438	-0.180	0.120
-2	0.189	-0.178	0.139	0.544	-0.262	0.098	-0.421	-0.167	0.095
-1	0.383	0.029	0.144	0.388	-0.139	-0.029	-0.150	0.154	-0.016
0	1.000	0.269	0.287	0.271	-0.341	-0.335	-0.068	-0.211	-0.074

оцененная модель для общей занятости выглядит следующим образом:

$$\Delta_4 y_t = y_t - y_{t-4} = 107.18 \Delta_4 \log(\text{ВВП})_t - 65.67 \Delta_4 \log(\text{INV})_{t-1} + \eta_t$$

$$\eta_t = -0.482 \eta_{t-1} + 7.320 + a_t$$

где зависимая переменная  $y_t$  – занятость в тыс. человек, ВВП (ВВП) и инвестиции (INV) – реальные значения в млн. сомов, взятые в логарифмической форме (поэтому сезонная разность приближенно равна темпу роста), и  $a_t$  – ошибки белого шума. Используя эту оцененную модель, мы прогнозируем занятость на период с 1-го кв. 2014 г. по 4-й кв. 2016 г.



**Рисунок15:** *Остатки подогнанной модели для ряда занятости*

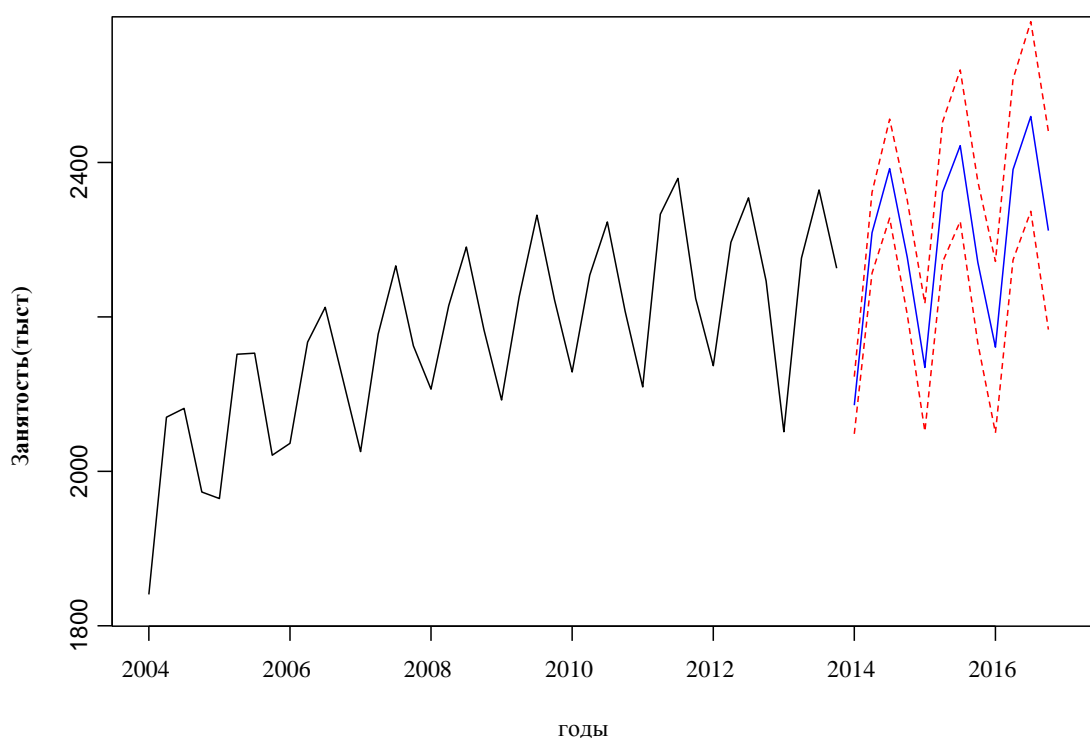
Для первого периода прогнозирования (1 кв. 2014– 4 кв. 2014) значения ВВП и инвестиций имеются в наличии. Для остального периода прогнозирования с 1-го кв. 2015 по 4-й кв. 2016 прогнозы будут возможны, только если мы спрогнозируем будущие значения изменений ВВП и инвестиций. Прогноз по ВВП был получен из МЭ. Реальные значения были получены с использованием прогнозированных дефляторов, а также из МЭ. Прогноз по инвестициям основан на сезонных моделях ARIMA, класс которых, описан в Разделе 2. Выбранная модель для инвестиций - ARIMA (0,0,1)(0,1,0)<sub>[4]</sub> с дрейфом, и ARIMA (0,1,0)(0,1,1) без дрейфа, соответственно, для ВВП и зарплаты. Значения прогноза с 95-процентными интервалами предсказания представлены в Таблице 6 и на Рисунке 16.

Важно отметить, что интервалы предсказания в этой модели не принимают во внимание неопределенность в прогнозах предикторов, и поэтому они, вероятно, занижены.

**Таблица6:**Прогнозные значения занятости(тыс.), полученные путем регрессирования квартальных изменений занятости по квартальным изменениям ВВП и квартальным изменениям инвестиций с лагом, с членами ошибок  $AR(1)$  и 95-процентными интервалами предсказаний. Период: 1-кв.2014– 4-й кв. 2016 г.

Квартал	Прогноз	Ниж.95	Верх.95
2014 кв.1	2083.4	2009.3	2157.5
2014 кв.2	2309.4	2204.7	2414.2
2014 кв.3	2403.1	2274.8	2531.4
2014 кв.4	2298.6	2150.4	2446.7
2015 кв.1	2098.7	1933.1	2264.3
2015 кв.2	2334.1	2152.7	2515.5
2015 кв.3	2429.7	2233.7	2625.6
2015 кв.4	2310.4	2100.9	2519.9
2016 кв.1	2126.7	1904.5	2348.9
2016 кв.2	2361.4	2127.2	2595.7
2016 кв.3	2452.5	2206.8	2698.2
2016 кв.4	2338.3	2081.7	2594.9

**Прогнозы общей занятости**



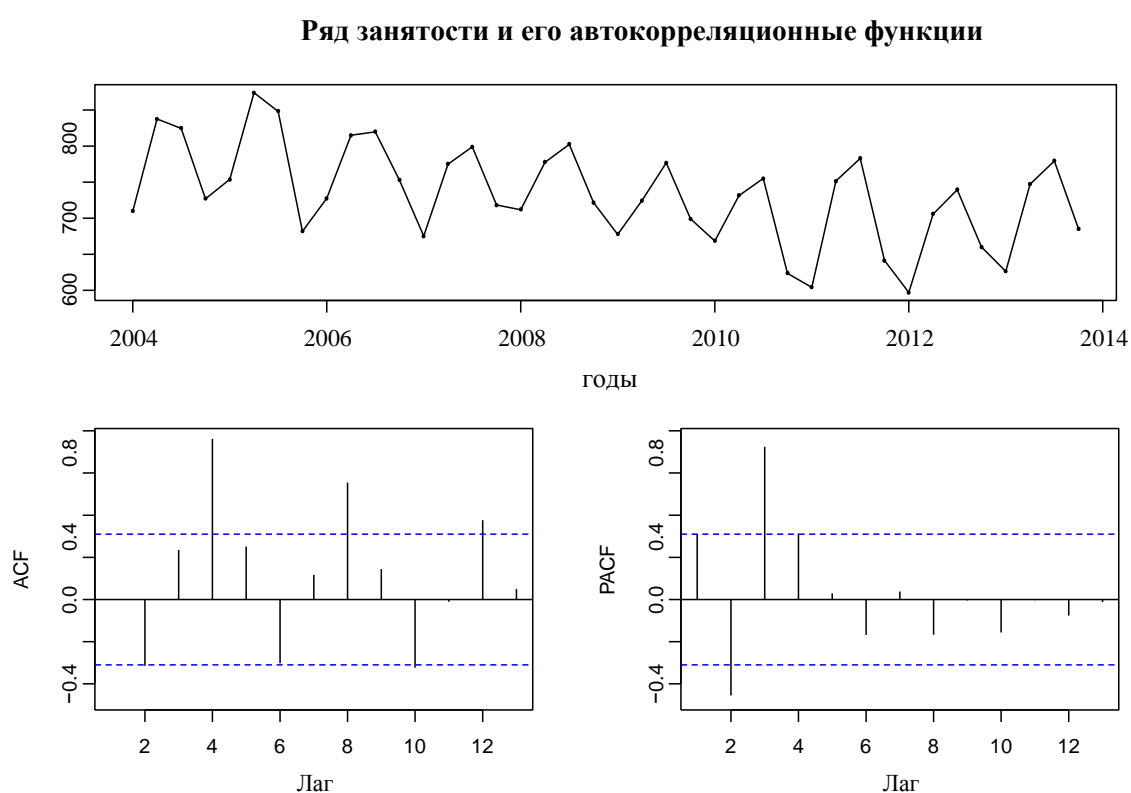
**Рисунок16:**Прогнозы, полученные путем регрессирования квартальных изменений занятости по квартальным изменениям ВВП и инвестиций с лагом, с использованием модели с ошибками  $AR(1)$ , с указанием 95-процентных интервалов предсказаний. Период: 1-кв. 2014 – 4-й кв. 2016 г.

### 3.2.3 Занятость по секторам деятельности

Аналогичным образом, мы прогнозируем занятость для четырех главных секторов –сельского хозяйства, промышленности, строительства и услуг.

#### Сельское хозяйство

Данные, очевидно, нестационарные с сильной сезонностью (см. рисунок 17), в связи с чем, мы сначала берем сезонную разность. Грубое разложение ряда занятости в сельском хозяйстве на тренд-цикл, сезонную и нерегулярную компоненту показано на рисунке 18. Данные после взятия сезонной разности (см. Рисунок 19) достаточно стационарные, поэтому дополнительные раунды взятия разностей не требуются.

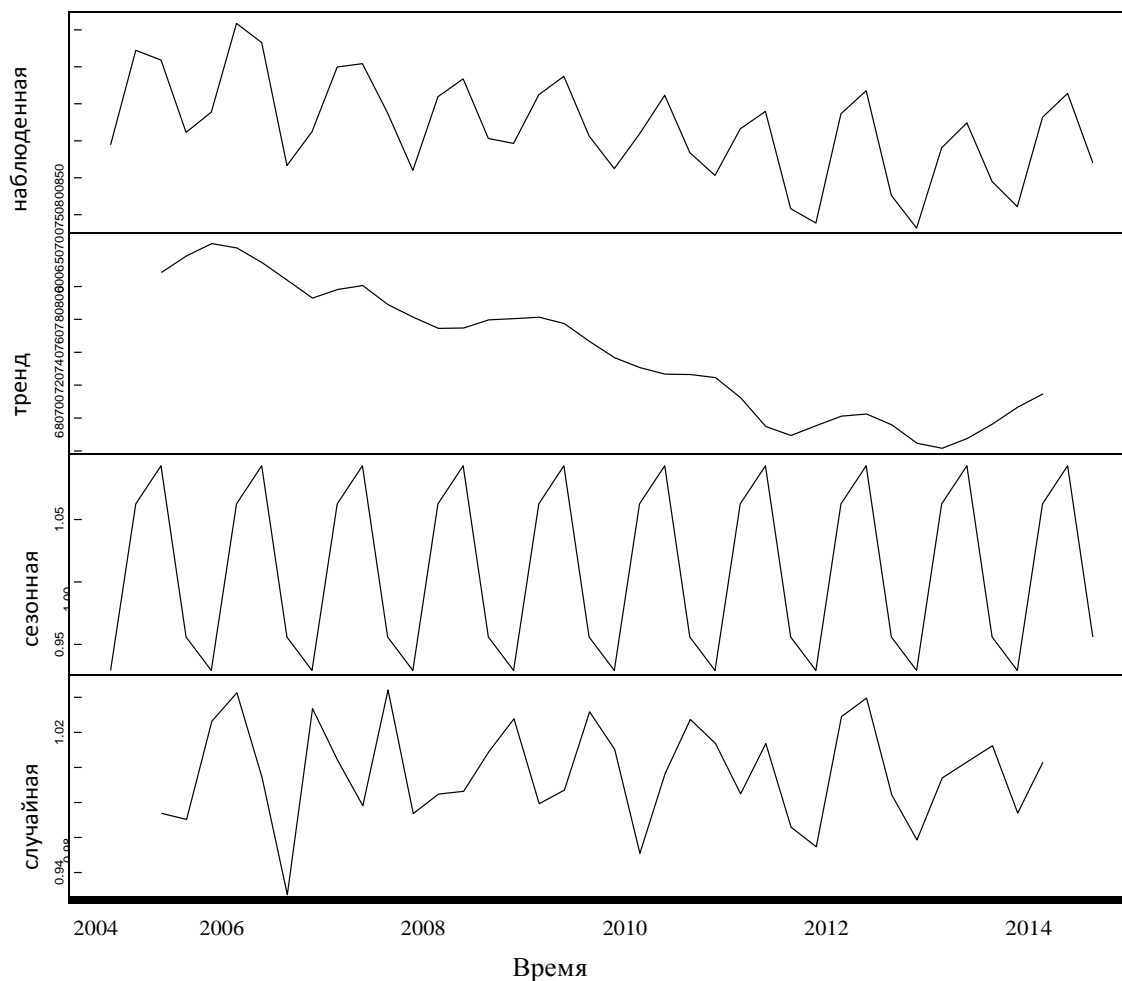


**Рисунок 17:** Ряд занятости в сельском хозяйстве и его автокорреляционные функции

Изучение структуры корреляции для квартальных изменений ряда показывает, что макроэкономические индикаторы не вносят вклада в объяснение его временного характера, за исключением ряда производительности труда<sup>18</sup>. Фактически, как показано в Таблице 7, все коэффициенты намного меньше 0,4, что свидетельствует о слабой и незначительной корреляции. В связи с этим, прогноз занятости в сельском хозяйстве основывается исключительно на внутренних свойствах ряда, и применяется сезонная модель ARIMA.

<sup>18</sup> Обратите внимание, что в этом случае мы также исключили из потенциальных предикторов ряд производительности.

### Разложение мультипликативного временного ряда



**Рисунок 18:** Мультипликативное разложение ряда занятости на его компоненты

Исходя из свойств временного ряда, мы выбрали ARIMA (0, 0,1)(0,1,1)<sub>[4]</sub> с дрейфом.

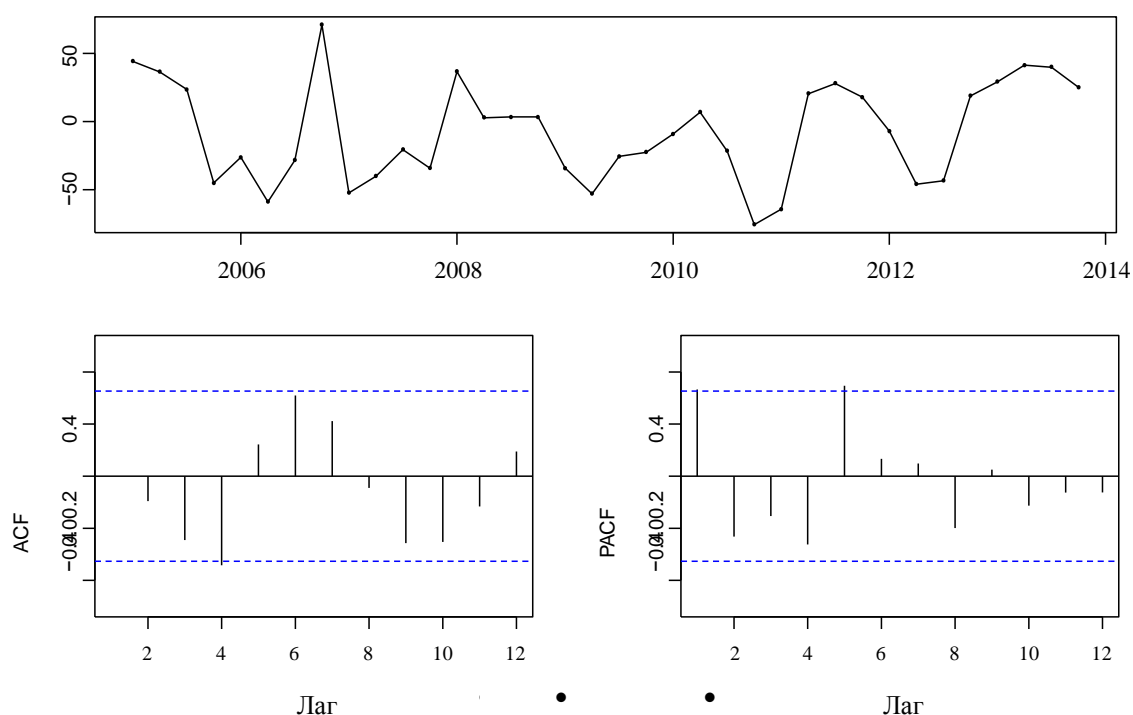
Оцененная сезонная модель ARIMA:

$$\Delta_4 y_t = -2.804 + a_t + 0.595a_{t-1} - 0.840a_{t-4}$$

которая проходит необходимые проверки. Остатки в подогнанной модели ARIMA представлены на рисунке 20, показывающем, что остатки, судя по всему, являются белым шумом. Прогнозы, полученные с помощью модели на последующие три года, показаны на рисунке 21.



### Занятость в сельском хозяйстве после взятия сезонной разности

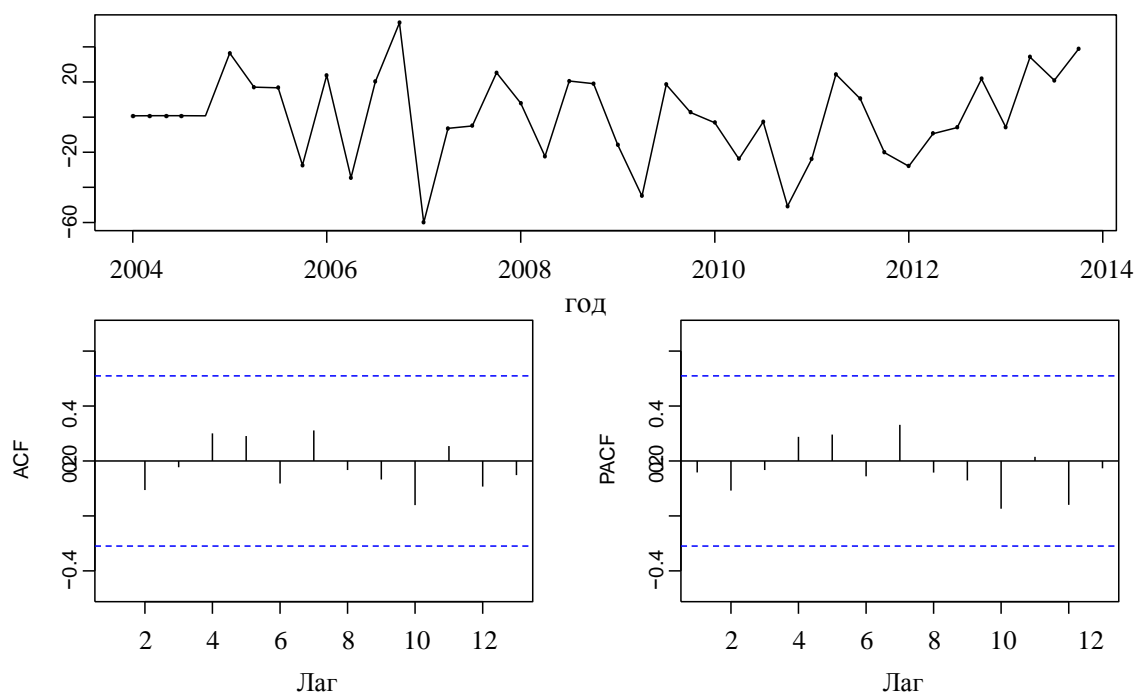


**Рисунок 19:** Занятость в сельском хозяйстве после взятия сезонной разности и ее автокорреляционные функции

**Таблица 7:** Структура кросс-корреляции квартальных изменений между занятостью в сельском хозяйстве и потенциальными макроэкономическими предикторами

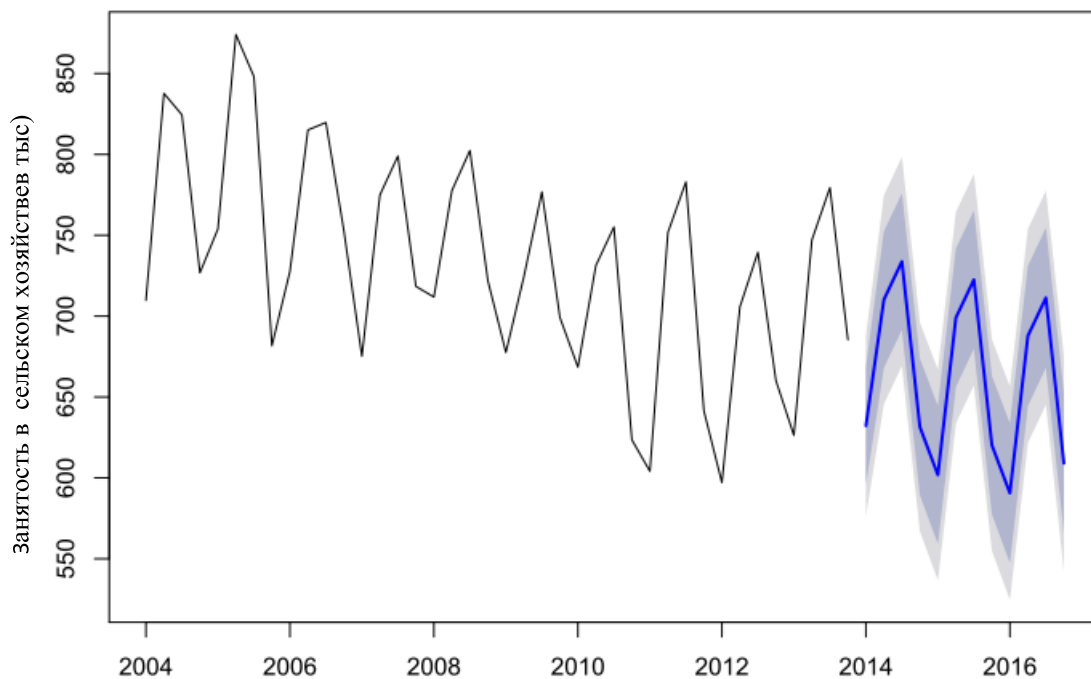
Лаг	Занятс/х	ВВП	ДСс/хг	ДСпрм	ДССтр	ДСУсл	Ч.потреб	Об. потреб	Инв
-4	-0.342	-0.043	0.132	-0.169	0.125	-0.015	-0.165	0.133	0.038
-3	-0.245	-0.218	-0.114	-0.245	0.123	0.08	0.031	-0.122	0.052
-2	-0.096	-0.229	-0.086	-0.213	0.037	0.084	-0.021	-0.295	-0.004
-1	0.332	0.036	0.043	0.094	0.183	0.067	0.007	-0.117	0.024
0	1	0.285	-0.084	0.357	0.052	0.075	0.259	-0.054	0.029
Лаг	Занятс/х	Эксп	Имп	Зарпл	Произ. всего	Произ. с/х	Произ. Прм.	Произ. стр	Произ. усл
-4	-0.342	-0.033	-0.037	-0.222	-0.05	0.305	-0.223	0.099	-0.155
-3	-0.245	-0.102	0.024	-0.015	-0.224	0.058	-0.252	0.046	-0.035
-2	-0.096	-0.16	-0.137	0.133	-0.137	-0.01	-0.168	0.078	0.073
-1	0.332	0.043	-0.051	0.043	0.117	-0.145	0.177	0.311	0.164
0	1	0.136	0.2	-0.176	0.091	-0.645	0.41	0.138	0.249

### Ошибки ARIMA



**Рисунок 20:** Остатки подогнанной модели  $ARIMA(0, 0, 1)(0, 1, 1)[4]$  с дрейфом для ряда занятости в сельском хозяйстве

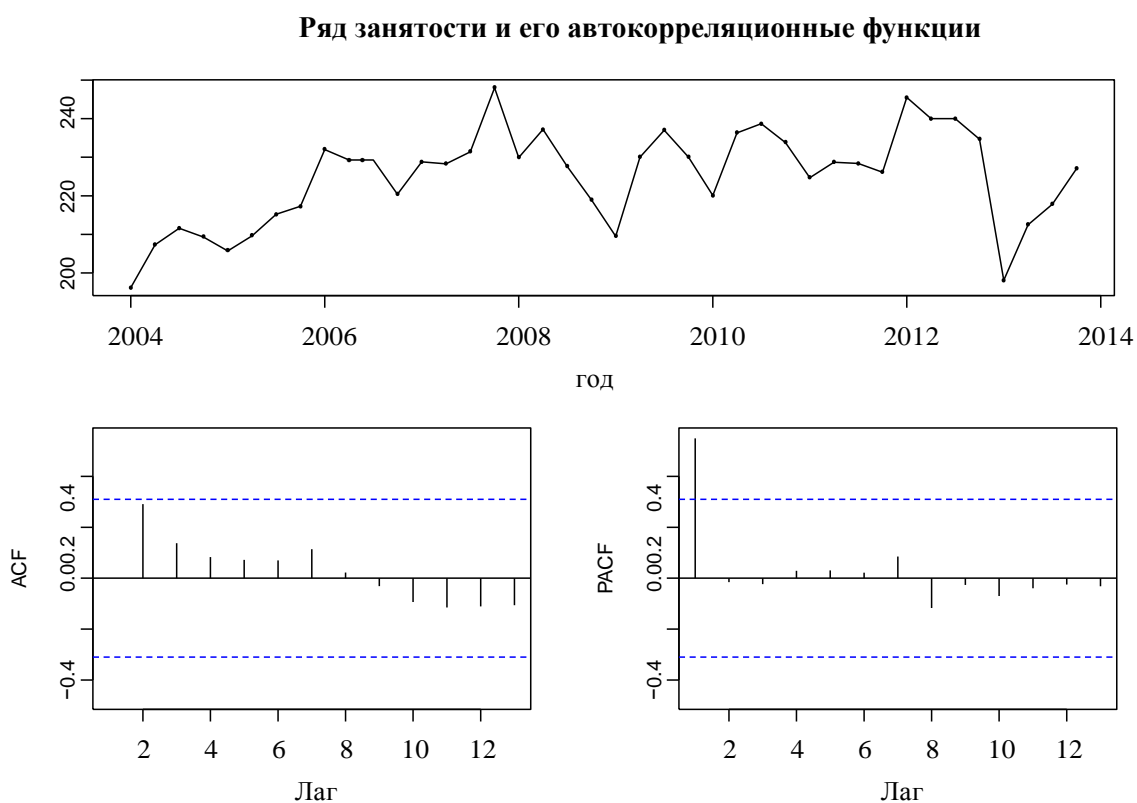
### Прогнозы занятости в сельском хозяйстве



**Рисунок 21:** Прогноз занятости в сельском хозяйстве с использованием  $ARIMA(0, 0, 1)(0, 1, 1)[4]$  с дрейфом, с указанием 80-процентного и 95-процентного интервалов предсказаний

## Промышленность

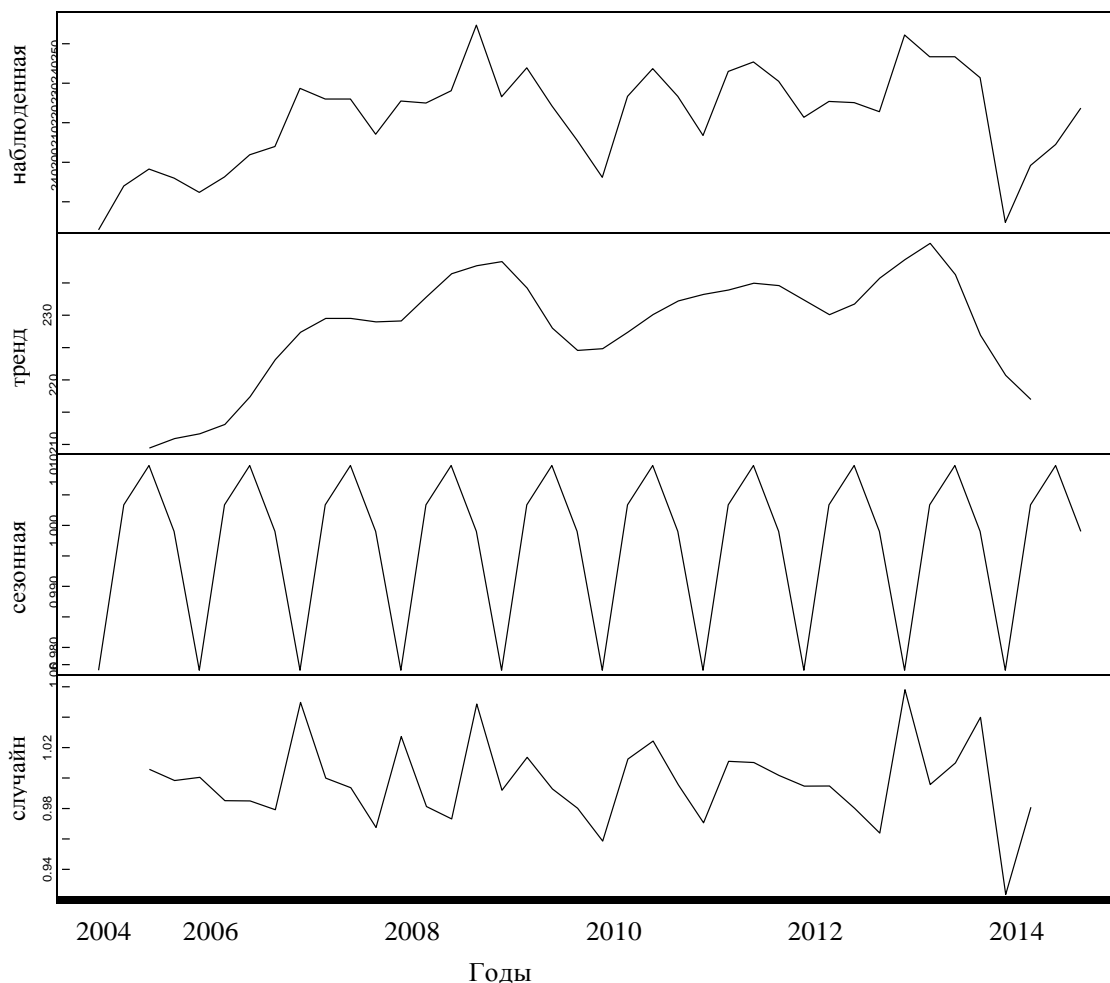
Ряд занятости в промышленности демонстрирует незначительную сезонность и довольно стабильный характер с течением времени (см. рисунок 22), в связи с чем, взятие разностей ряда не является строгим требованием. Однако характер сезонности (см. разложение ряда занятости в промышленности на тренд-цикл, сезонность и нерегулярную компоненту на рисунке 23) указывает на то, что перед рассмотрением потенциальных корреляций с предикторами следует убрать сезонные изменения. Данные после взятия сезонной разности и функции автокорреляции показаны на рисунке 24.



**Рисунок 22:** Ряд занятости в промышленности и его автокорреляционные функции

Структура корреляции между ежеквартальными изменениями ряда занятости в промышленной и некоторыми макроэкономическими индикаторами показывает, что некоторые коэффициенты корреляции больше, чем 0.4 (см. Таблицу 8): зарплата, добавленная стоимость, частное потребление, инвестиции.

### Разложение мультипликативного временного ряда



**Рисунок 23:** Мультипликативное разложение ряда занятости на его компоненты.

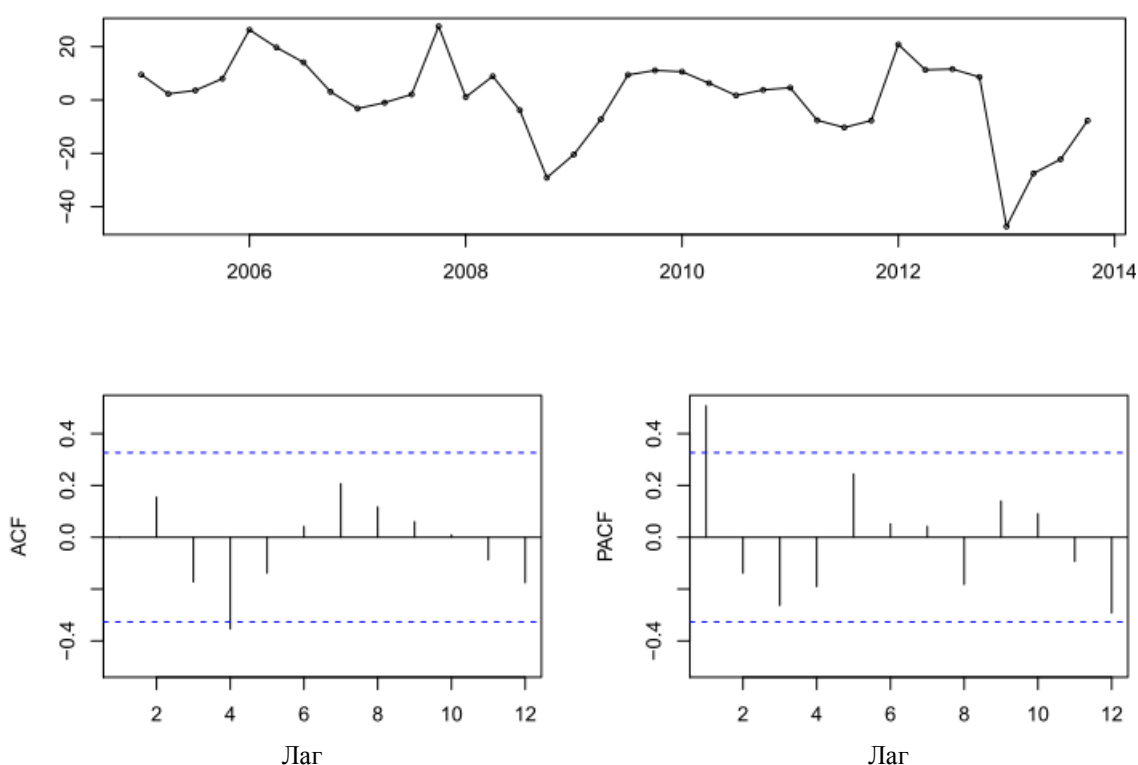
Процедура выбора переменной сохраняет в качестве предикторов добавленную стоимость и зарплату, однако остатки регрессии должны быть сформированы моделью ARIMA. Порядок модели ARIMA выбирается в соответствии со значением  $AIC_c$ , что приводит нас к ARIMA (1,0,0)(0,1,0)<sub>[4]</sub>. Оцененная модель для занятости в промышленности имеет следующий вид:

$$\Delta_4 y_t = -13.47 \Delta_4 \log(\text{ДС пром.})_t + 41.60 \Delta_4 \log(\text{Зарпл})_t + \eta_t$$

$$\eta_t = 0.414 \eta_{t-1} + a_t$$

Для первого периода прогнозирования (1- кв. 2014– 4-й кв. 2014 г.) значения зарплат и добавленной стоимости имеются в наличии. Для остального периода с 1-го кв. 2015 по 4-й кв. 2016 г. прогноз занятости в промышленности требует прогнозных будущих значений квартальных изменений зарплат.

### Занятость в промышленности после взятия сезонной разности



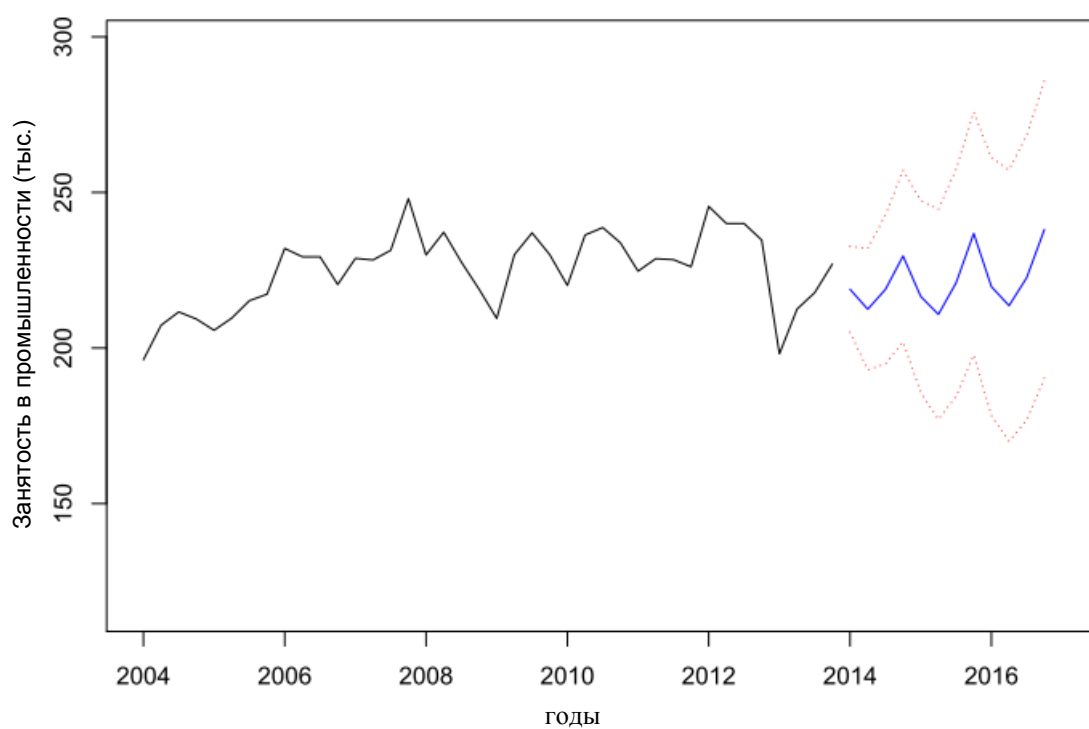
**Рисунок24:** Занятость в промышленности после взятия сезонной разности и ее автокорреляционные функции.

Прогноз по зарплате основан на модели  $ARIMA(0,1,0)(0,1,1)$  с дрейфом. Прогнозные значения добавленной стоимости в промышленности на период с 1-го кв. 2015 по 4-й кв. 2016 г. были предоставлены МЭ. Прогнозы, полученные с помощью модели, на последующие три года показаны на рисунке 25.

**Таблица 8:** Структура кросс-корреляции квартальных изменений между занятостью в промышленности и потенциальными макроэкономическими предикторами.

Лег	Занятость	ВВП	ДС С/х	ДС прм	ДС Стр	ДС Усл	Ч. потреб	Об. потреб	Инв
-4	-0.35	0.10	-0.20	0.07	0.07	0.32	0.16	-0.11	-0.01
-3	-0.17	-0.08	0	-0.06	-0.06	0.28	0.27	0.05	0.14
-2	0.15	-0.04	0.02	-0.10	-0.10	0.22	0.37	0.15	0.26
-1	0.51	-0.13	-0.09	-0.20	-0.02	0.08	0.37	0.09	0.4
0	1.00	-0.33	-0.09	-0.37	-0.03	-0.01	0.06	0.17	0.31
Лег	Зан. прм	Эксп	Имп	Зарпл	Прон. об	Пронз. с/х	Пронз. Прм.	Пронз. Стр.	Пронз. усл
-4	-0.35	0.24	0.21	0.06	0.16	-0.17	0.15	-0.02	0.36
-3	-0.17	-0.01	0.33	0.1	0.06	-0.04	-0.01	-0.17	0.34
-2	0.15	-0.17	0.38	0.18	0.05	0.03	-0.13	-0.19	0.23
-1	0.51	-0.34	0.31	0.3	-0.15	0.04	-0.3	-0.12	-0.05
0	1	-0.20	0.13	0.47	-0.47	0.17	-0.59	-0.12	-0.30

### Прогноз занятости в промышленности

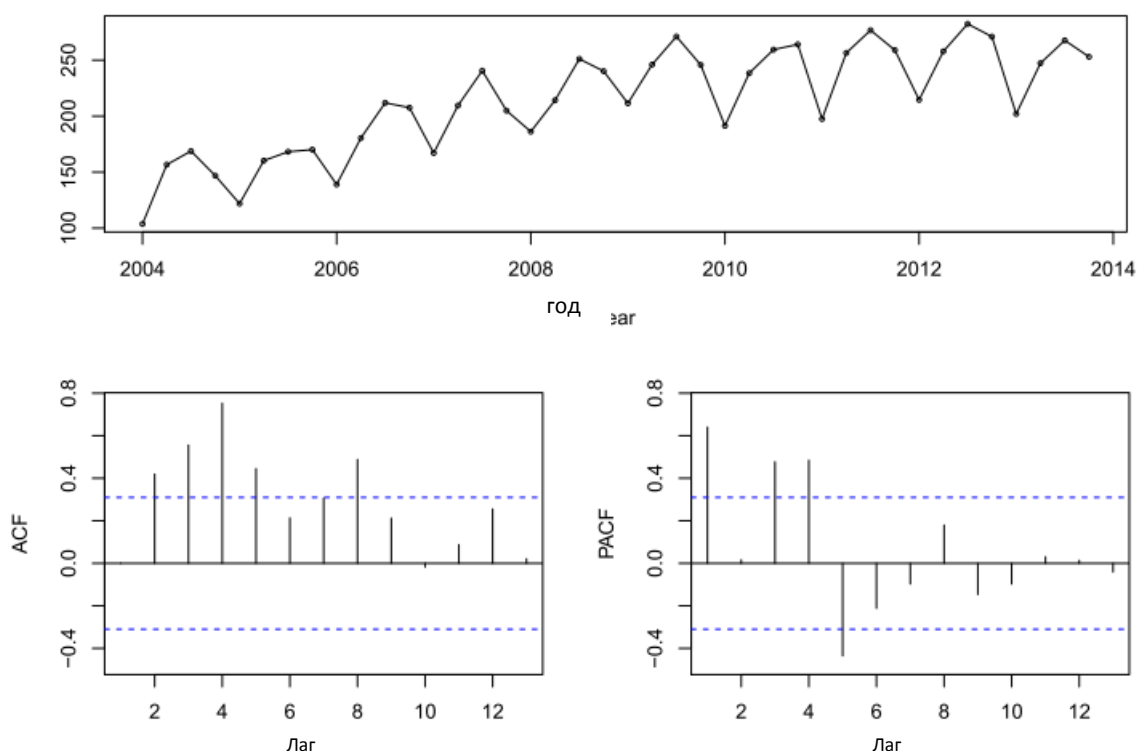


**Рисунок 25:** Прогноз занятости в промышленности, как функция текущей добавленной стоимости сектора и зарплат; с ошибками  $ARIMA(1, 0, 0)(0, 1, 0)$  и с указанием 95-процентного интервала предсказаний

## Строительство

Графики ряда и его автокорреляционные функции (см. Рисунок 26) демонстрируют сильную сезонность, указывая на то, что по ряду строительства необходимо взять, по крайней мере, сезонную разность, чтобы получить стационарность. После взятия сезонной разности, ряд (см. Рисунок 28) представляется достаточно стационарным, поэтому дополнительные раунды взятия разностей не проводились.

**Ряд занятости и его автокорреляционные функции**



**Рисунок 26:** Ряд занятости в строительстве и его автокорреляционные функции.

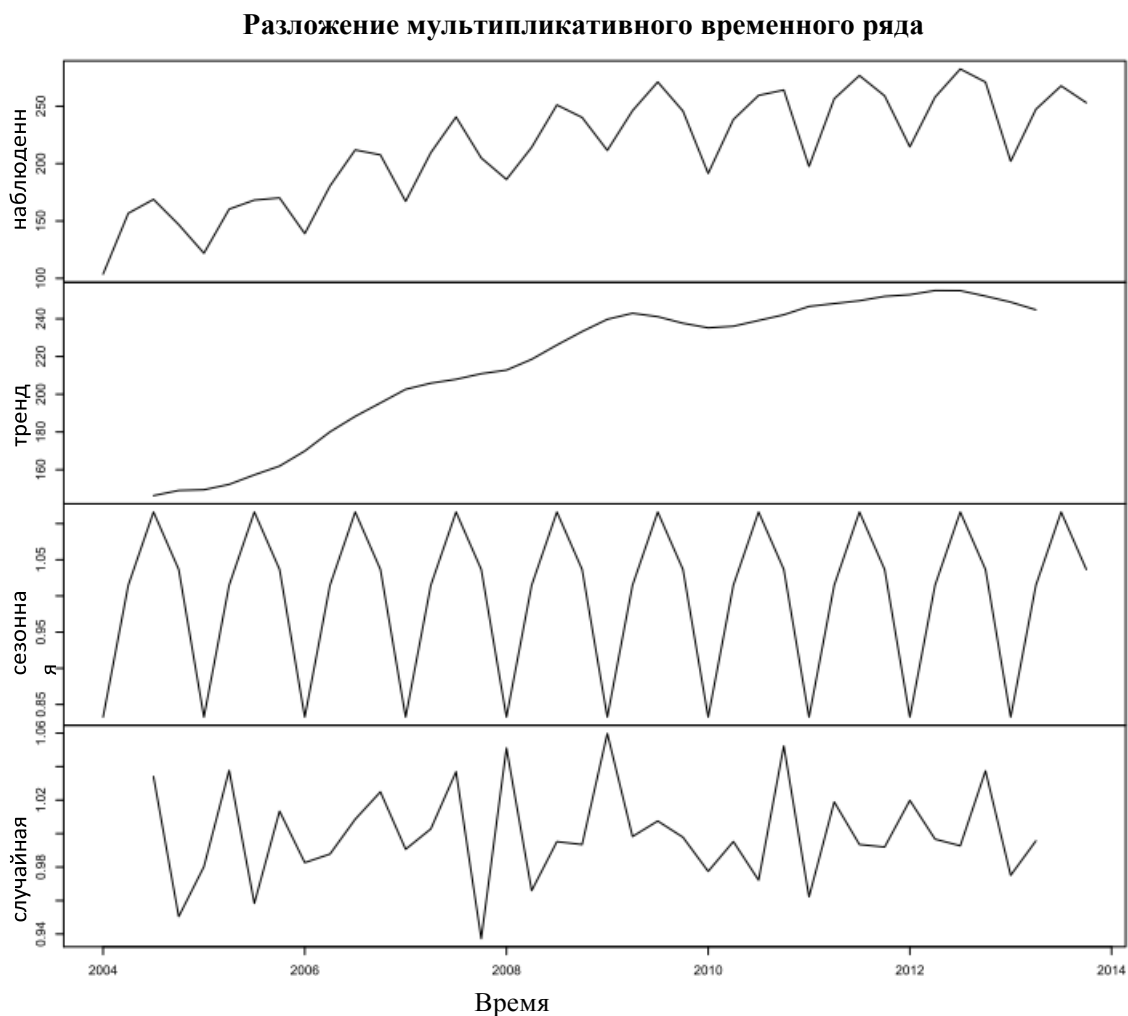
Структура корреляции, представленная в Таблице 9, показывает, что потенциальным предиктором является зарплата с различными лагами и в меньшей степени, импорт и частное потребление.

В подогнанной регрессионной динамической модели единственным значимым предиктором является текущая зарплата. Поэтому мы производим регрессию квартальных изменений занятости в строительстве только по квартальным изменениям зарплаты, обеспечивая при этом соответствующую структуру для ошибок ARIMA. Выбранная модель ARIMA, которая включает эндогенные стационарные регрессоры и ошибки ARMA, представляет собой ARIMA (2,0,0)(1,1,0), которая выглядит следующим образом:

$$\Delta_4 y_t = 68.36 \Delta_4 \log(\text{Зарпл})_t + \eta_t$$

$$\eta_t = 0.359 \eta_{t-1} + 0.452 \eta_{t-2} - 0.593 \eta_{t-4} + a_t$$

и проходит требуемые проверки остатков.

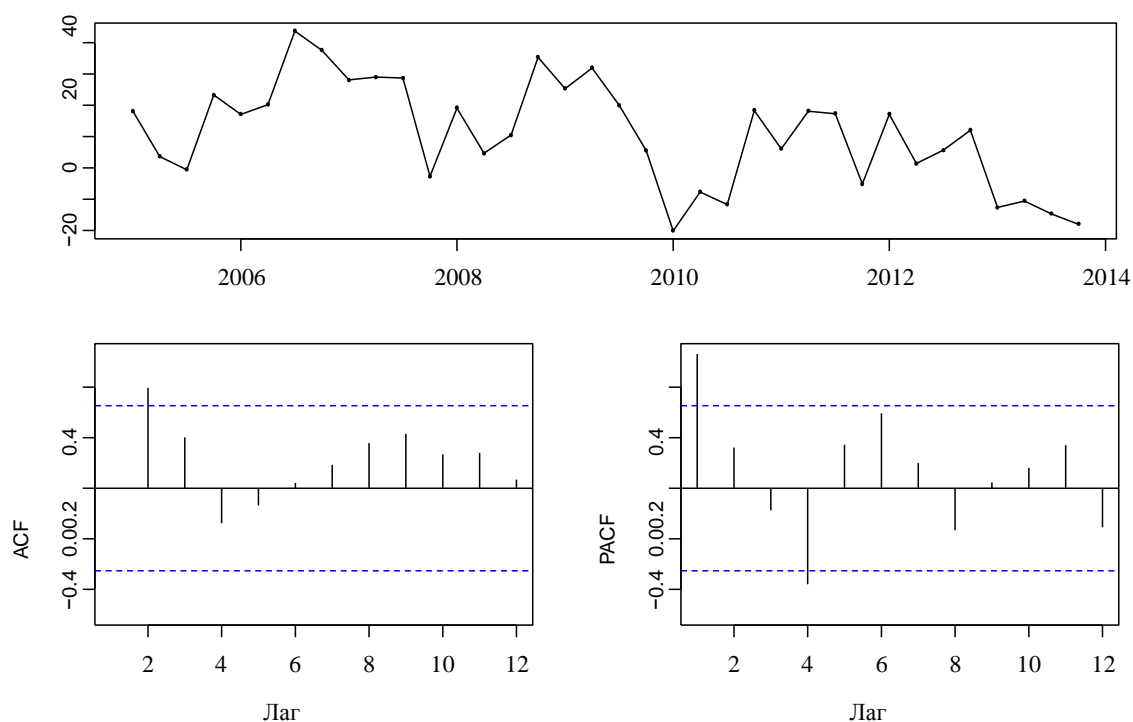


**Рисунок 27:** Мультипликативное разложение ряда занятости на его компоненты

Прогноз для строительного сектора возможен только в случае, если имеются значения будущих изменений в зарплате. Значения переменной «зарплаты» имеются в наличии для первого года прогноза, но нам необходимо прогнозировать их для следующих восьми кварталов (1-й кв. 2015 г. –4-й кв. 2016 г.). Наилучшей моделью ARIMA для прогнозирования зарплаты является ARIMA (0,1,0)(0,1,1) с дрейфом.



### Занятость в строительстве после взятия сезонной разности



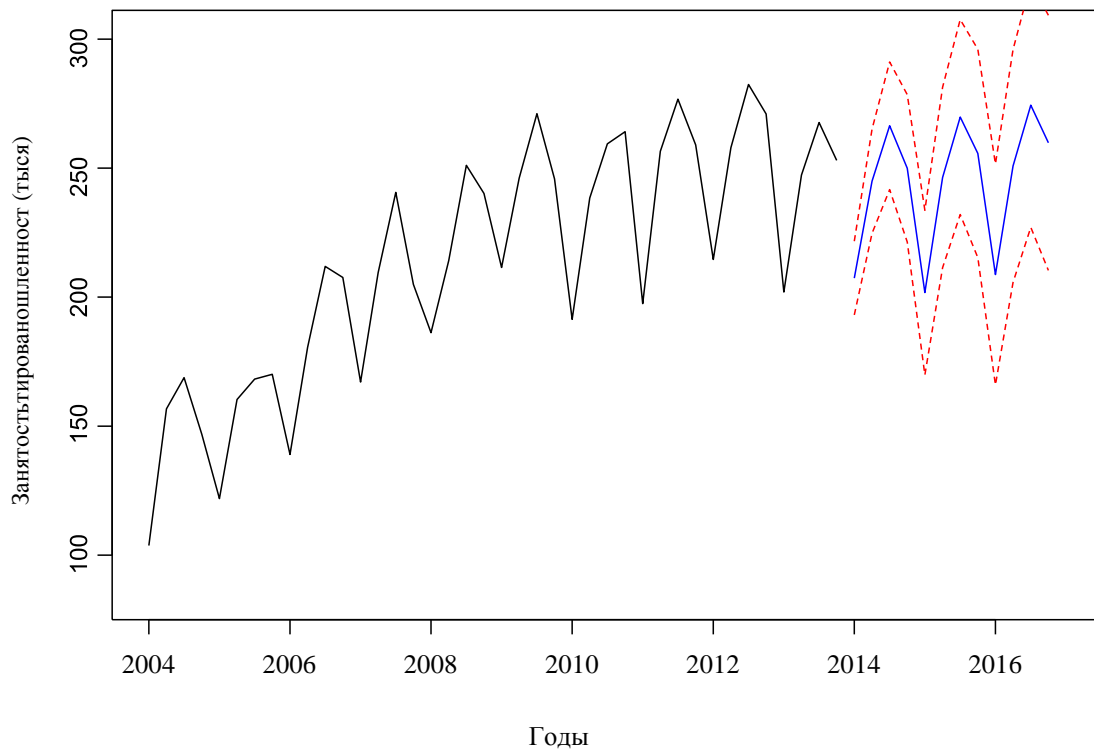
**Рисунок 28:**Ряд занятости в строительстве после взятия сезонной разности и его автокорреляционные функции.

С прогнозируемыми значениями зарплаты можно предсказать занятость в строительстве в течение всего периода с 1-го кв. 2014 г. по 4-й кв. 2016 г., ряд которой показан на Рисунке29.

**Таблица 9:**Структура кросс-корреляции квартальных изменений между занятостью в строительстве и потенциальными макроэкономическими предикторами.

Лег	Занят стр.	ВВП	ДС С/х	ДС прм	ДС Стр	ДС Усл	Ч. потреб	Об. потреб	Инв
-4	0.067	-0.292	-0.077	-0.301	-0.027	0.135	0.171	-0.061	0.248
-3	0.223	-0.35	-0.047	-0.309	-0.265	0.069	0.231	0.045	0.172
-2	0.435	-0.281	-0.014	-0.226	-0.275	-0.011	0.049	0.278	0.179
-1	0.705	-0.109	-0.102	-0.032	-0.213	-0.079	-0.045	0.222	0.092
0	1.000	-0.049	-0.017	0.015	-0.24	-0.02	-0.21	-0.057	-0.225
Лег	Занят стр.	Эксп	Имп	Зарпл	Проз. всего	Проз. с/х	Проз. Прм.	Проз. стр	Проз. Усл.
-4	0.067	-0.073	0.275	0.329	-0.223	0.083	-0.313	-0.206	0.075
-3	0.223	-0.136	0.292	0.431	-0.318	0.118	-0.367	-0.436	-0.057
-2	0.435	-0.128	0.174	0.32	-0.335	0.094	-0.315	-0.393	-0.248
-1	0.705	-0.007	0.013	0.213	-0.347	0.023	-0.180	-0.316	-0.438
0	1.000	0.189	-0.149	0.173	-0.301	0.205	-0.134	-0.347	-0.541

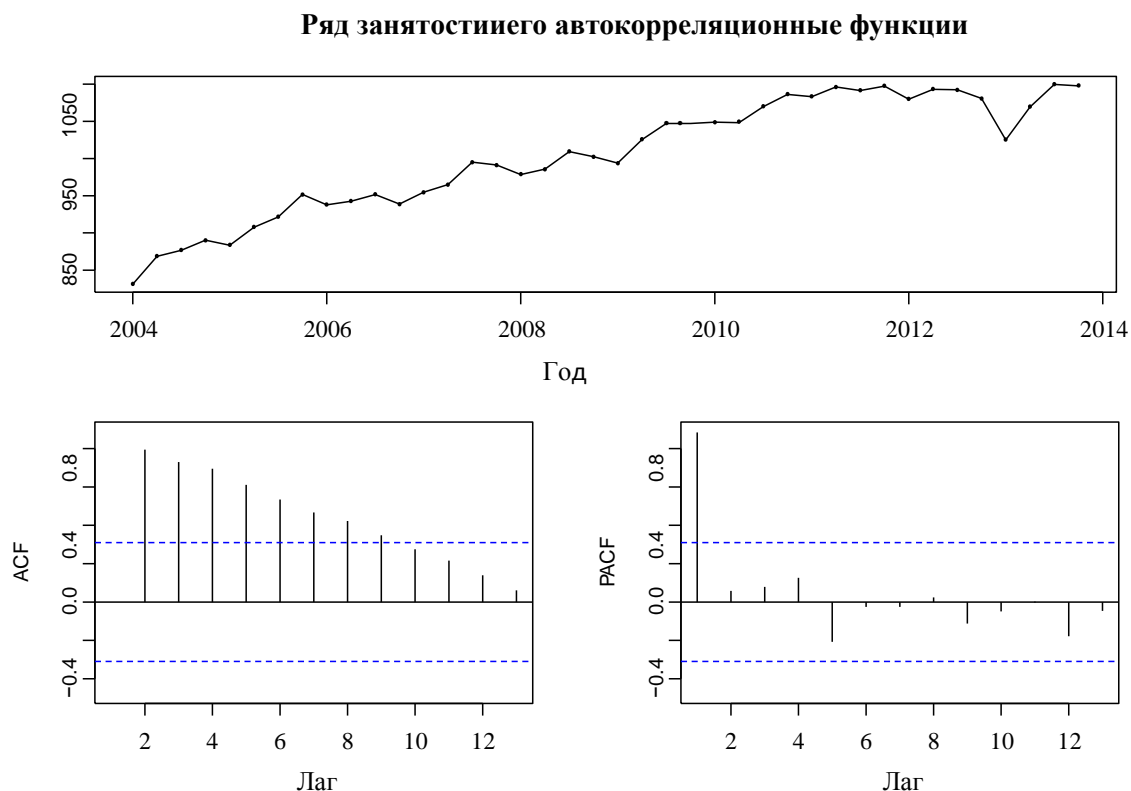
### Прогнозы занятости в строительстве



**Рисунок 29:** Прогноз занятости в строительстве, как функция текущей зарплаты; с ошибками  $ARIMA(2, 0, 0)(1, 1, 0)$ , с указанием 95-процентных интервалов предсказаний.

## Услуги

Ряд занятости в услугах не показывает сильной сезонности, однако, очевидно не является стационарным (см. Рисунок 30). Грубое разложение ряда на его компоненты представлено на Рисунке 31. Обратите внимание на то, что оцененные сезонные коэффициенты очень близки к единице.



**Рисунок 30:** Ряд занятости в услугах и его автокорреляционные функции

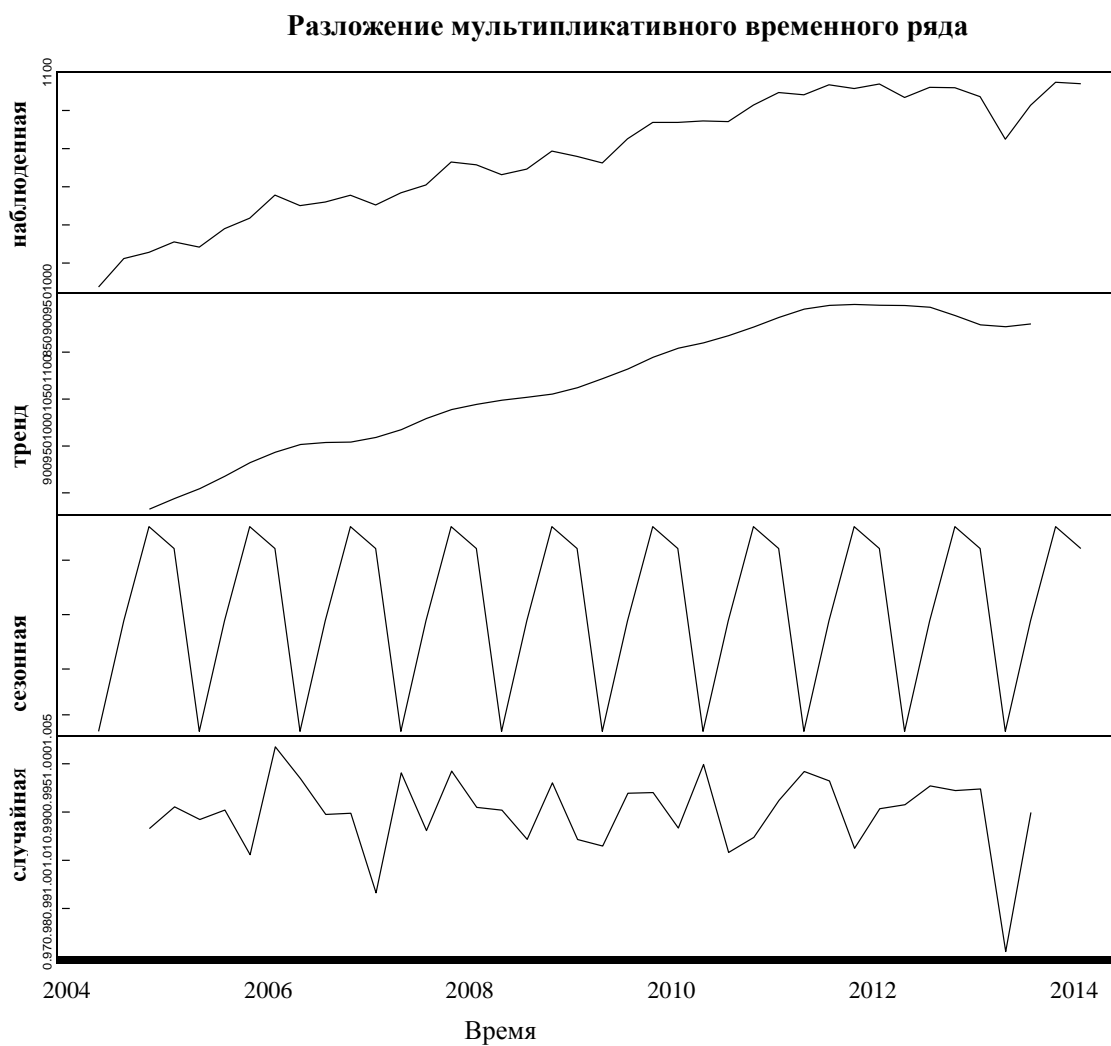
Таким образом, мы берем первые разности (т.е., разности с лагом 1), чтобы получить стационарные данные, показанные на Рисунке 32.

Структура кросс-корреляции, приведенная в Таблице 10, указывает на следующие переменные, как потенциальные предикторы изменений между последовательными кварталами в оригинальном ряду: текущий ВВП, добавленная стоимость в сельском хозяйстве, добавленная стоимость в промышленности с лагом на один период, добавленная стоимость в услугах, частное потребление, инвестиции, экспорт и импорт.

Дальнейшее изучение динамических регрессионных моделей, вдохновленное принципом экономичности, привело к включению в качестве потенциальных предикторов только текущего значения ВВП и значений добавленной стоимости в промышленности с лагом. Оценка следующей динамической регрессионной модели:

$$\Delta_1 y_t = 5.512 + 21.951 \Delta_1 \log(\text{ВВП})_t - 19.632 \Delta_1 \log(\text{ДС Прм})_{t-1} + a_t$$

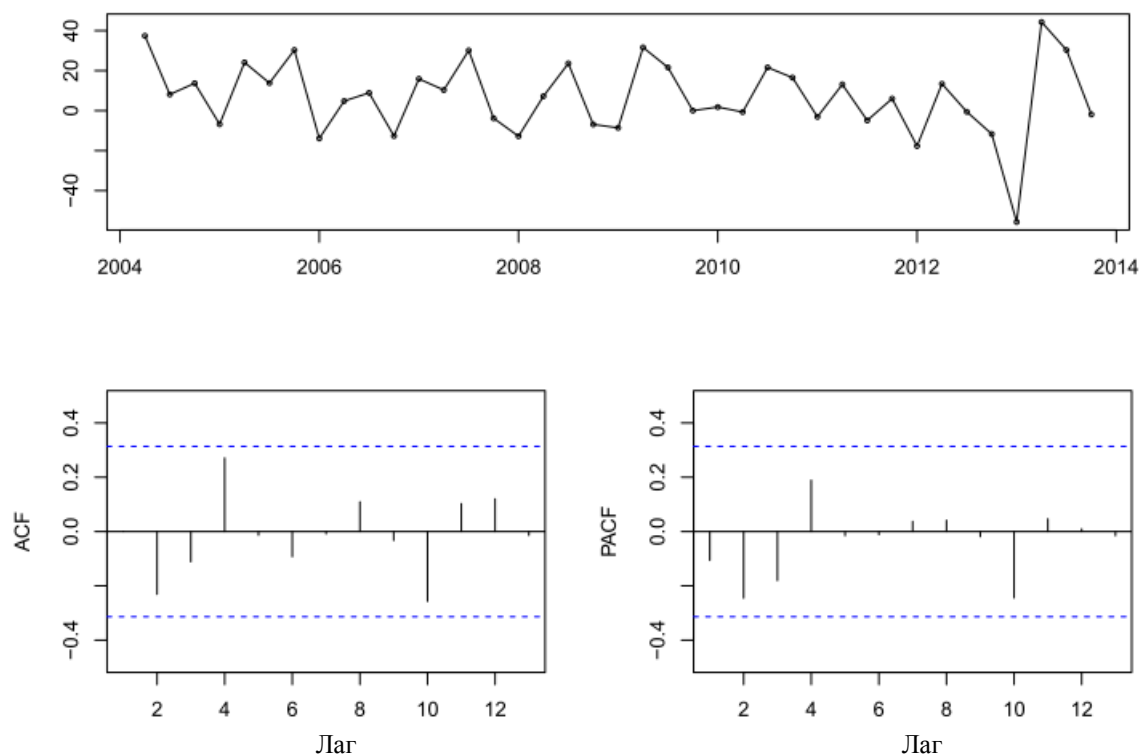
Привела к хорошей подгонке и остаткам, которые в виде ряда не коррелируются, устраняя потребность в формировании ошибочных членов с помощью модели ARIMA (см. Рисунок 33).



**Рисунок 31:** Мультипликативное разложение ряда занятости на его компоненты

Для получения прогноза занятости в услугах, нам необходимы будущие значения ВВП и добавленной стоимости в промышленности для периода с 1 кв. 2015 г. по 4-й кв. 2016 г. Эти значения могут быть получены из таблиц прогнозирования МЭ. Полученный в результате прогнозный ряд представлен на Рисунке 34.

### Занятость в секторе услуг после взятия разностей

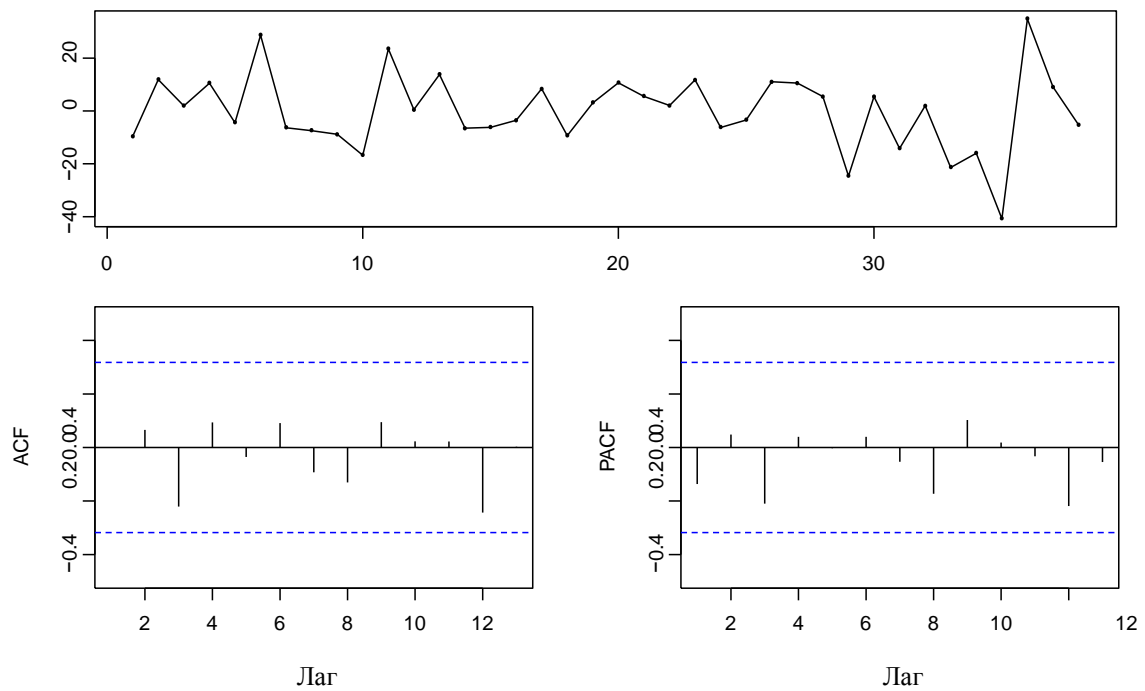


**Рисунок 32:** *Ряд занятости в услугах после взятия первой разности и его автокорреляционные функции.*

**Таблица 10:** *Структура кросс-корреляции квартальных изменений между занятостью в услугах и потенциальными макроэкономическими предикторами.*

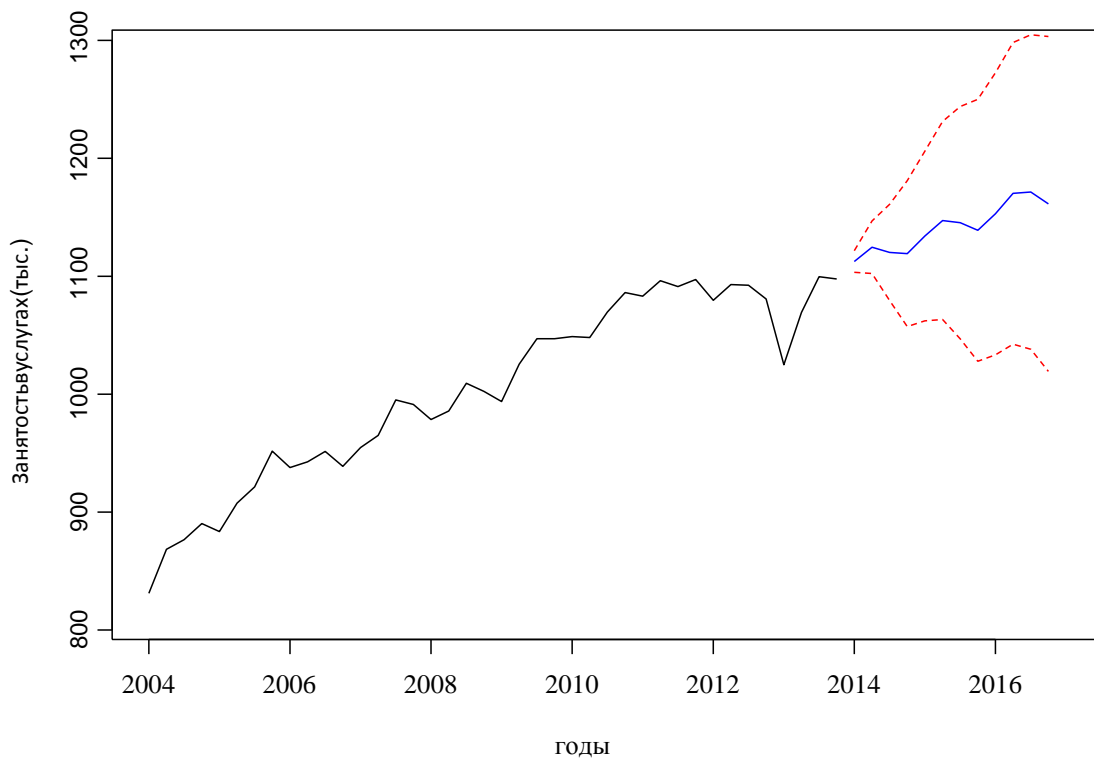
Лег	Занятусл.	ВВП	ДС с/х	ДСпрм	ДССтр	ДСУсл	Ч. потреб	Об. потреб.	Инвест
-4	0.271	0.353	0.344	-0.134	0.384	0.314	0.342	0.281	0.397
-3	-0.111	-0.148	-0.008	-0.510	-0.164	-0.356	-0.12	-0.276	-0.190
-2	-0.231	-0.465	-0.518	0.076	-0.268	-0.129	-0.475	-0.065	-0.243
-1	-0.106	0.148	0.088	0.515	-0.056	0.108	0.212	-0.005	0.016
0	1.000	0.580	0.532	-0.029	0.553	0.504	0.515	0.456	0.488
Лег	Занятусл.	Эксп	Имп	Зарпл	Произ. всего	Произ. с/х	Произ. прм	Произ.стр	Произ.усл
-4	0.271	0.015	0.249	0.256	0.311	0.331	-0.143	0.345	0.304
-3	-0.111	-0.493	-0.311	-0.237	-0.235	-0.054	-0.502	-0.208	-0.36
-2	-0.231	0.155	-0.117	0.005	-0.401	-0.507	0.113	-0.24	-0.118
-1	-0.106	0.179	0.14	0.015	0.231	0.140	0.476	-0.009	0.120
0	1.000	0.323	0.432	0.399	0.505	0.522	-0.139	0.493	0.451

### Остатки регрессионной модели для услуг



**Рисунок 33:** Остатки регрессионной модели для прогнозирования занятости в услугах и ее автокорреляционные функции

### Прогнозы занятости в услугах



**Рисунок 34:** Прогноз занятости в услугах, как функция текущего ВВП и добавленной стоимости в промышленности с лагом, с указанием 95-процентных интервалов предсказаний.

## Выверка

Сумма рядов занятости по секторам должна быть равна ряду общей занятости. Так как мы проводили прогнозирование отдельно, необходимо согласовать *по факту* прогнозные значения. Сохраняя фиксированными прогнозы общей занятости, мы распределили несоответствия между прогнозом общей занятости и агрегатом секторных прогнозов пропорционально по каждому сектору. Эти несоответствия, несмотря на скромную величину, не являются пренебрежимо малыми, и составляют от 0,52% в 1-м квартале 2016 г. до к 3,25% в 3-м квартале 2016 г. Имеются другие варианты, которые могут быть далее обсуждены. Таблица 11 показывает прогноз занятости в четырех секторах после коррекции агрегированной величины.

**Таблица 11:** Прогноз общей занятости и по секторам деятельности

	Сельское хоз.	Промышл.	Строительство	Услуги	Общая
2014кв.1	607.38	210.25	199.23	1068.79	2085.64
2014кв.2	715.44	214.02	246.72	1132.88	2309.06
2014кв.3	750.35	223.76	272.44	1145.54	2392.08
2014кв.4	644.78	234.40	255.19	1142.80	2277.17
2015кв.1	596.31	214.58	199.93	1123.76	2134.58
2015кв.2	716.74	216.15	252.66	1176.36	2361.91
2015кв.3	741.90	226.84	277.02	1176.12	2421.88
2015кв.4	625.48	238.81	257.85	1148.58	2270.73
2016кв.1	587.53	218.57	207.69	1147.04	2160.84
2016кв.2	708.06	219.90	258.22	1204.77	2390.94
2016кв.3	735.21	230.10	283.60	1210.74	2459.64
2016кв.4	620.73	242.67	264.79	1183.66	2311.85

### 3.3 Сезонная коррекция

Мы, наконец, можем заняться потенциальной сезонностью в квартальных временных рядах общей занятости и занятости по секторам деятельности (см. Таблицу 11), а также в квартальном ряде серии безработицы (см. Таблицу 4). Придерживаясь международных методов наиболее успешной практики мы скорректируем квартальные ряды занятости и безработицы, применяя модель X-13ARIMA-SEATS, разработанную и поддерживаемую Бюро переписи населения США<sup>19</sup>. Эта модель представляет недавнюю модернизацию и улучшение предыдущей версии X-11, с некоторыми улучшениями, направленными непосредственно на методы преодоления проблем коррекции, и, тем самым, на расширение диапазона экономических временных рядов, которые могут адекватно корректироваться с учетом сезонных изменений. Она объединяет и расширяет возможности более старых моделей X-12ARIMA (разработанной Бюро переписи населения США) и TRAMO-SEATS (разработанной Банком Испании).

Ряды, скорректированные с учетом сезонных изменений, представлены в Таблице 12 и на Рисунках 35, 36 и 37.<sup>20</sup>

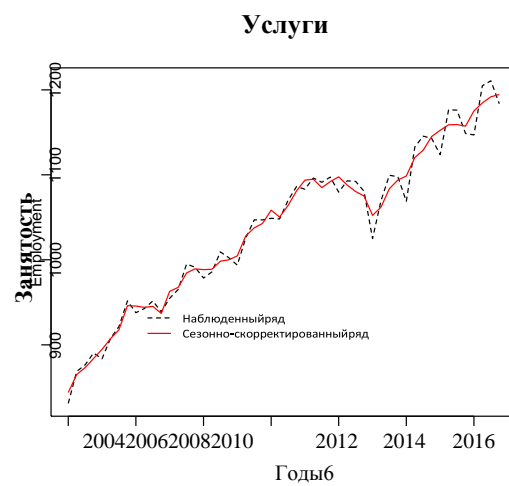
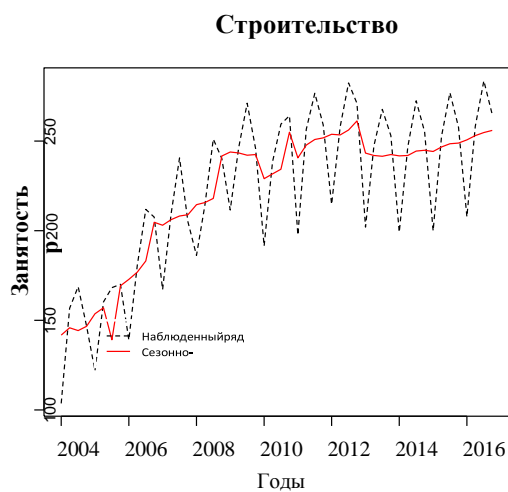
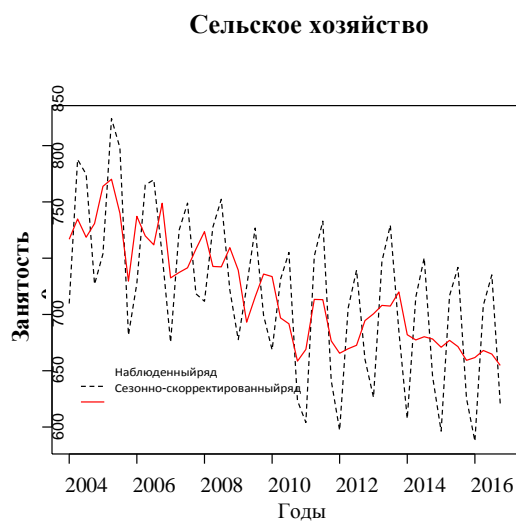
**Таблица 12:** Прогноз занятости по секторам деятельности, общей занятости и безработицы – данные скорректированы с учетом сезонных изменений

	Сельское хоз.	Промышл.	Строительст.	Услуги	Занятость	Безработ.
2014кв.1	681.97	219.30	241.74	1098.68	2241.70	212.1
2014кв.2	677.42	217.14	241.90	1120.49	2256.94	212.1
2014кв.3	680.17	220.01	244.40	1129.10	2273.68	210.6
2014кв.4	678.26	223.68	244.82	1145.04	2291.81	211.7
2015кв.1	670.87	223.77	244.13	1152.41	2291.18	214.1
2015кв.2	677.05	220.22	246.67	1158.92	2302.86	215.0
2015кв.3	671.51	223.46	248.50	1159.14	2302.60	214.2
2015кв.4	659.29	226.74	248.87	1157.25	2292.15	215.0
2016кв.1	661.76	223.21	250.66	1175.19	2310.82	216.2
2016кв.2	668.02	224.50	252.99	1184.69	2330.18	217.9
2016кв.3	664.81	227.15	254.71	1191.89	2338.56	217.6
2016кв.4	654.76	237.81	255.92	1194.58	2343.07	218.2

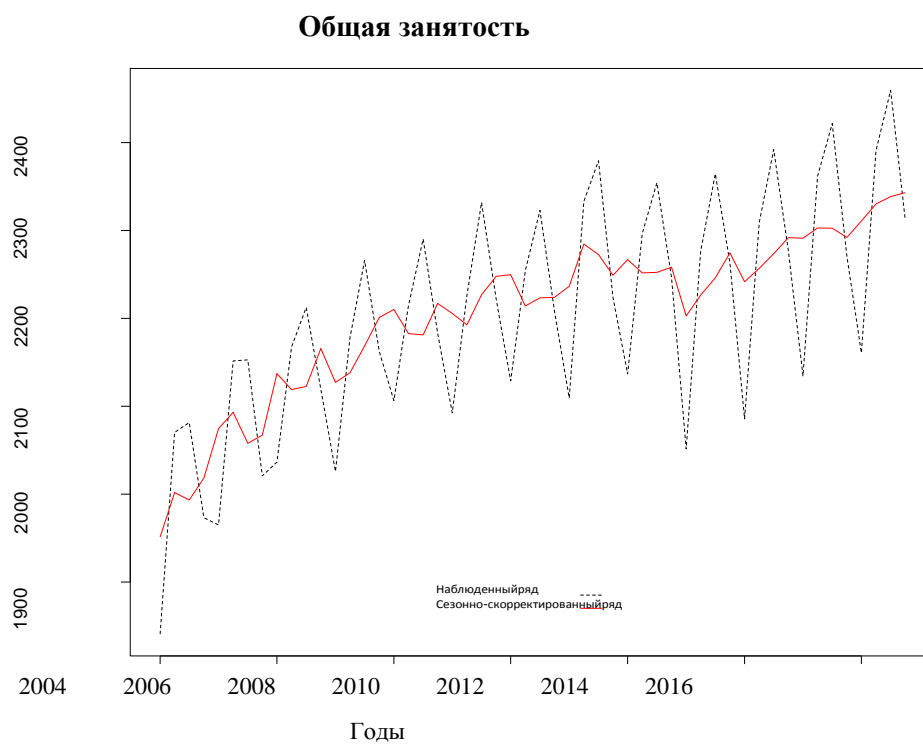
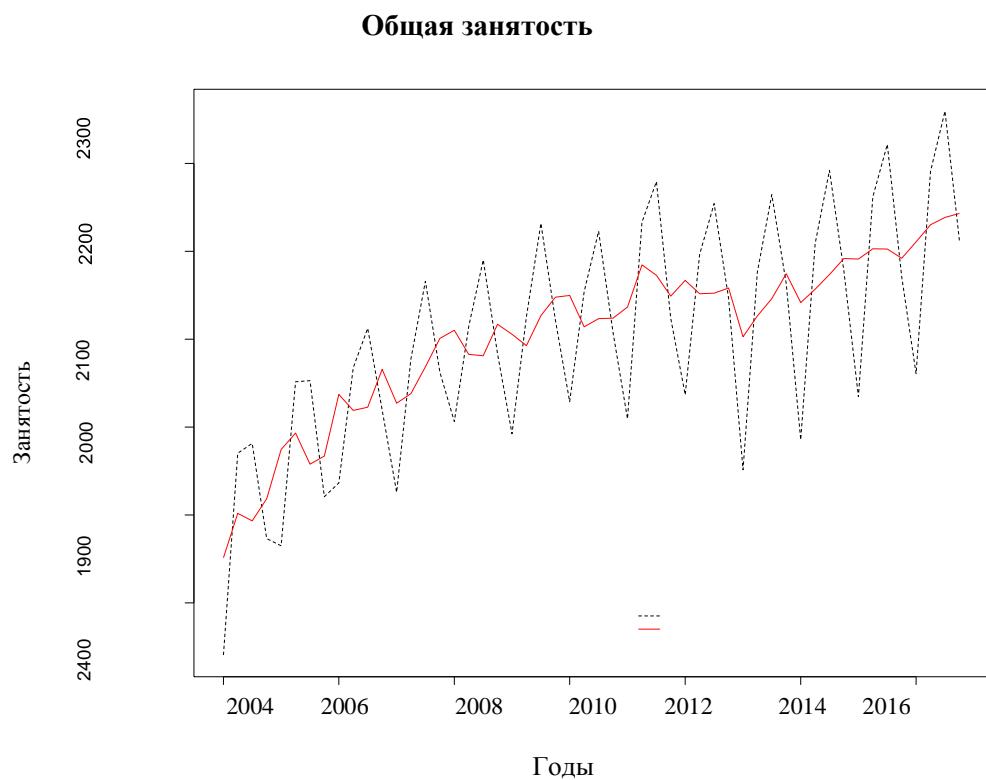
<sup>19</sup> Технические подробности см. в Справочном руководстве по X13-ARIMA-SEATS, Версия 1.1 Бюро переписи США (US Census Bureau (2015), "X13-ARIMA-SEATS Reference Manual. Version 1.1"), размещенном на <http://www.census.gov/srd/www/x13as/>

<sup>20</sup> Все вычисления в рамках X13-ARIMA-SEATS были выполнены с использованием финальной функции в сезонном пакете, имеющемся в сети R Archivenetwork (RD evelopmentCoreTeam, 2014).

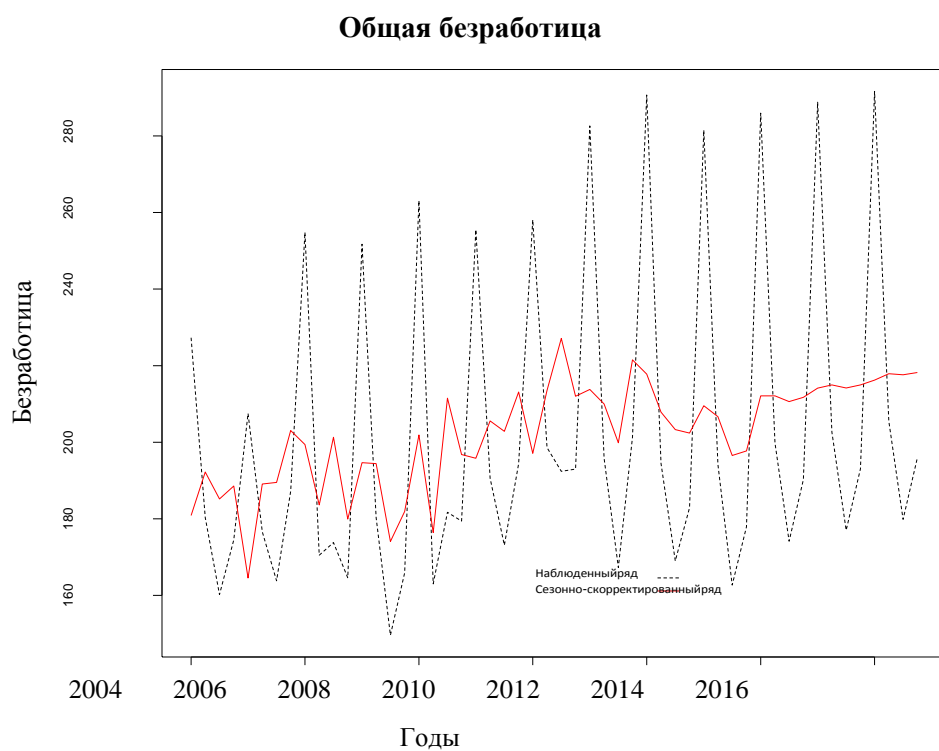




**Рисунок 35:** Прогноз занятости, скорректированный на сезонные изменения по секторам деятельности.



**Рисунок 36:** Прогноз занятости, скорректированной на сезонные изменения



**Рисунок 37:** Прогноз безработицы, скорректированной на сезонные изменения

