**ВСЕРОССИЙСКАЯ АКАДЕМИЯ ВНЕШНЕЙ ТОРГОВЛИ**

**Факультет международного бизнеса**

# Предметно-аналитическая справка по эконометрике

на тему: «Исследование стоимости фиксированного набора потребительских товаров и услуг»

Работу выполнил:
Студент 1 курса ФМБ
Варламов Иван

Научный руководитель:
доктор физмат наук,
профессор Д.В. Филиппов

Работа сдана на кафедру
«  »_____ 2019г.

Оценка работы
_____
«  »_____ 2019г.

**Москва 2019**

## 1. Подбор данных для анализа

Цель проводимого исследования – построение зависимости стоимости фиксированного набора потребительских товаров и услуг от следующих параметров:

– Курс национальной валюты по отношению к доллару США;

– Доллар к евро обменный курс;

– Мировая цена на уран;

– Экспорт сырой нефти из Российской Федерации;

– Экспорт нефтепродуктов из Российской Федерации;

– Экспорт газа природного в газообразном состоянии из Российской Федерации;

– Цена на сырую нефть: Западный Техас, Средний уровень (WTI);

– Экспорт электроэнергии из Российской Федерации;

– Количество сотрудников: Горное дело и лесозаготовка: добыча нефти и газа.

### 1.1 Временной ряд №1 – XXX

XXX – стоимость фиксированного набора потребительских товаров и услуг; данные представлены в абсолютных значениях в рублях [рубль]. Источник: https://fedstat.ru/indicator/31052.

### 1.2 Временной ряд №2 – RU

RU – National Currency to US Dollar Exchange Rate: Average of Daily Rates for the Russia Federation (CCUSMA02RUM618N) – Курс национальной валюты по отношению к доллару США: средний дневных курсов для Российской федерации. Данные представлены в [рубль/$]. Источник: https://fred.stlouisfed.org/series/CCUSMA02RUM618N

### 1.3 Временной ряд №3 – EUR

EUR – U.S. / Euro Foreign Exchange Rate (EXUSEU) – Доллар к евро обменный курс. Данные представлены в [$/€]. Источник: https://fred.stlouisfed.org/series/EXUSEU.

### 1.4 Временной ряд №4 – PURAN

PURAN – Global price of Uranium (PURANUSDM) – Мировая цена на уран, данные представлены в абсолютных значениях в долларах за фунт [$/lb]. Источник: https://fred.stlouisfed.org/series/PURANUSDM.

### 1.5 Временной ряд №5 – EXOIL

EXOIL – Экспорт сырой нефти из Российской Федерации. Данные представлены в абсолютных значениях в тысячах долларов [$10^3$ $]. Источник: https://fedstat.ru/indicator/37393.

### 1.6 Временной ряд №6 – PCRUD

PCRUD – Экспорт нефтепродуктов из Российской Федерации. Данные представлены в абсолютных значениях в тысячах долларов [$10^3$ $]. Источник: https://fedstat.ru/indicator/37393.

### 1.7 Временной ряд №7 – PGAS

PGAS – Global price of Natural gas, EU (PNGASEUUSDM) – цена на природный газ. Данные представлены в абсолютных значениях в долларах за м$^3$ [$/м$^3$]. Источник: https://fred.stlouisfed.org/series/PNGASEUUSDM.

### 1.8 Временной ряд №8 – POIL

POIL – Crude Oil Prices: West Texas Intermediate (WTI) - Cushing, Oklahoma (MCOILWTICO) – Цены на сырую нефть: Западный Техас, Средний уровень (WTI) - Кушинг, Оклахома. Данные представлены в абсолютных значениях в долларах за баррель [$/bbls]. Источник: https://fred.stlouisfed.org/series/MCOILWTICO

### 1.9 Временной ряд №9 – EENER

PGAS – Экспорт электроэнергии из Российской Федерации. Данные представлены в абсолютных значениях в тысячах долларов [$10^3$ $]. Источник: https://fedstat.ru/indicator/37393.

### 1.10 Временной ряд №10 – EMP

EMP – All Employees: Mining and Logging: Oil and Gas Extraction (CES1021100001) – Все сотрудники: Горное дело и лесозаготовка: добыча нефти и газа. Данные в абсолютных значениях в тысячах человек [$10^3$ человек]. Данные для США.
Источник: https://fred.stlouisfed.org/series/CES1021100001.

## 1.11 Описательная статистика

Таблица 1.1 – Описательная статистика 1

| Временной ряд | Вариабельность | Среднее | Медиана | Минимум | Максимум |
|---|---|---|---|---|---|
| XXX | 0.29798 | 10728 | 10237 | 5421,2 | 16122 |
| RU | 0.38679 | 42,113 | 32,628 | 23,351 | 77,217 |
| EUR | 0.10217 | 1,2782 | 1,2993 | 1,0545 | 1,5759 |
| PURAN | 0.48840 | 44,390 | 40,778 | 18,568 | 136,22 |
| EXOIL | 0.29831 | $1{,}1064\cdot10^{7}$ | $1{,}0705\cdot10^{7}$ | $4{,}4275\cdot10^{6}$ | $1{,}7893\cdot10^{7}$ |
| PCRUD | 0.33221 | $6{,}4090\cdot10^{6}$ | $6{,}0286\cdot10^{6}$ | $2{,}7849\cdot10^{6}$ | $1{,}2550\cdot10^{7}$ |
| PGAS | 0.32641 | 8,8052 | 8,5600 | 3,4830 | 16,020 |
| POIL | 0.30886 | 74,239 | 74,120 | 30,320 | 133,88 |
| EENER | 0.32514 | 73717 | 69077 | 30419 | $1{,}3863\cdot10^{5}$ |
| EMP | 0.11516 | 168,32 | 162,60 | 139,70 | 200,80 |

Таблица 1.2 – Описательная статистика 2

| Времен-ной ряд | Ст. откл. | Ассиметрия | Эксцесс | 5% | 95% | IQ |
|---|---|---|---|---|---|---|
| XXX | 3196,6 | 0,090832 | -1,2879 | 5789,2 | 15578 | 5741,3 |
| RU | 16,289 | 0,54502 | -1,4056 | 24,490 | 66,981 | 29,786 |
| EUR | 0,13060 | 0,078805 | -0,87420 | 1,0808 | 1,4856 | 0,22380 |
| PURAN | 21,680 | 1,7800 | 4,0397 | 20,679 | 90,387 | 23,489 |
| EXOIL | $3{,}3004\cdot10^{6}$ | 0,13955 | -1,0847 | $6{,}3433\cdot10^{6}$ | $1{,}6608\cdot10^{7}$ | $5{,}8565\cdot10^{6}$ |
| PCRUD | $2{,}1291e\cdot10^{6}$ | 0,46486 | -0,44625 | $3{,}4277\cdot10^{6}$ | $1{,}0024\cdot10^{7}$ | $3{,}4357\cdot10^{6}$ |
| PGAS | 2,8742 | 0,24654 | -0,37266 | 4,0700 | 14,360 | 4,3325 |
| POIL | 22,929 | 0,19546 | -0,84884 | 41,474 | 106,40 | 41,040 |
| EENER | 23968 | 0,73107 | 0,10778 | 40121 | $1{,}2259\cdot10^{5}$ | 27674 |
| EMP | 19,383 | 0,24159 | -1,3959 | 142,10 | 198,10 | 36,600 |

## 1.12 Временные ряды

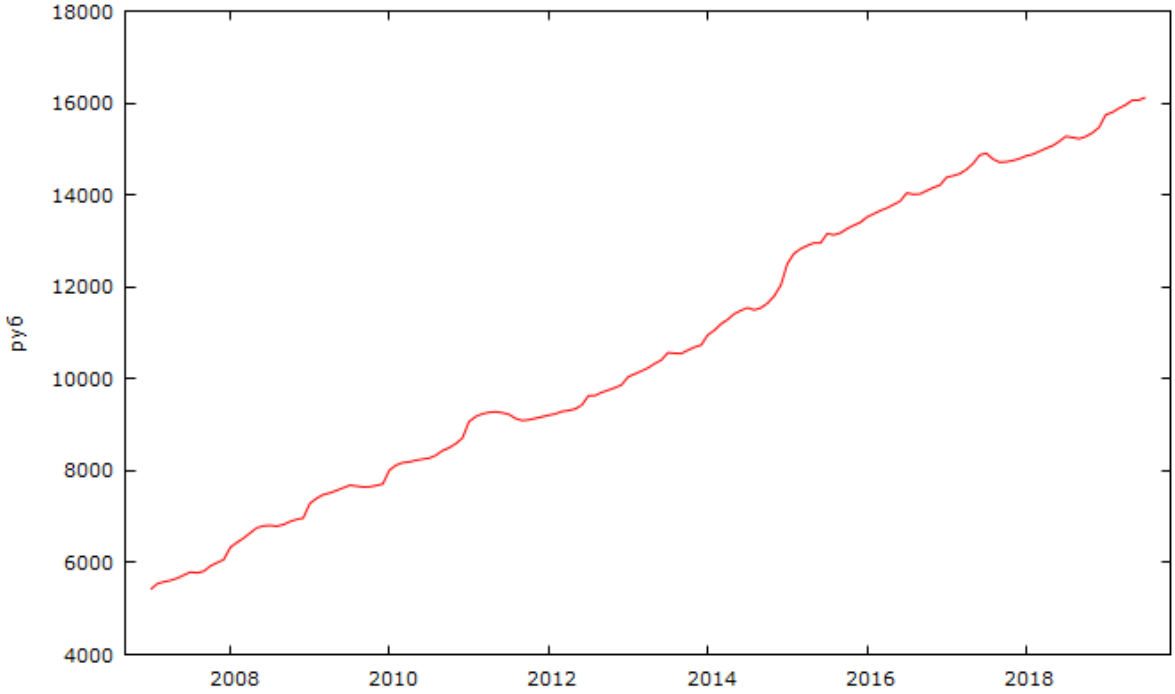



Рисунок 1.1 – Временной ряд XXX

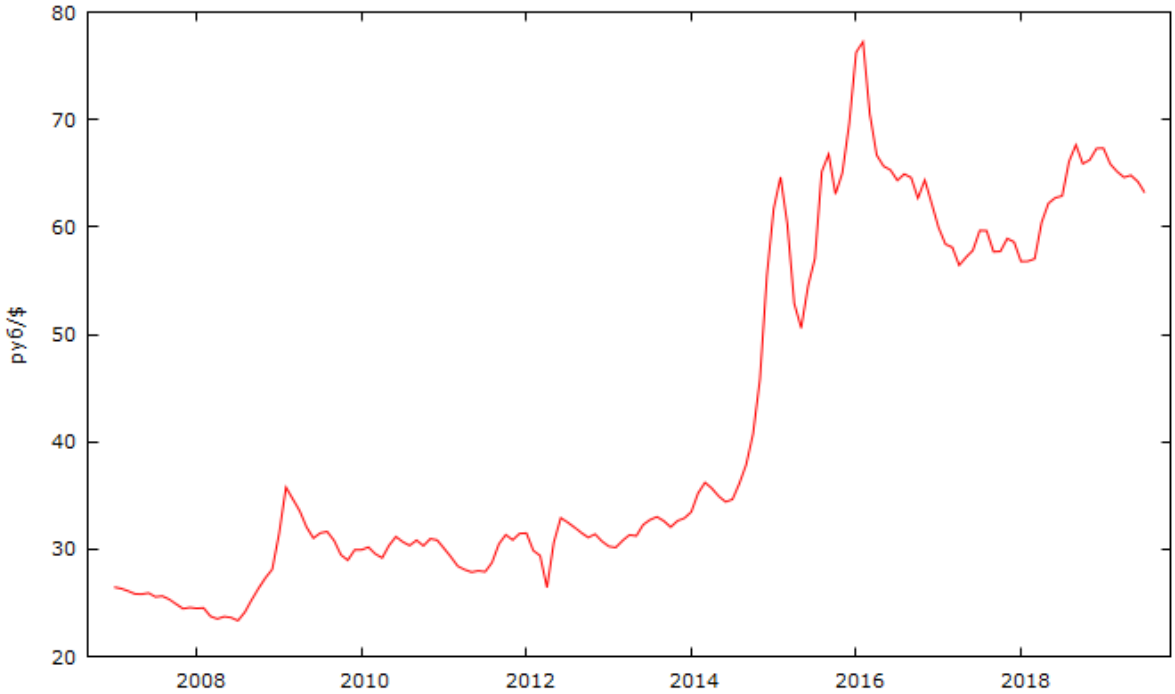



Рисунок 1.2 – Временной ряд RU

Рисунок 1.3 – Временной ряд EUR

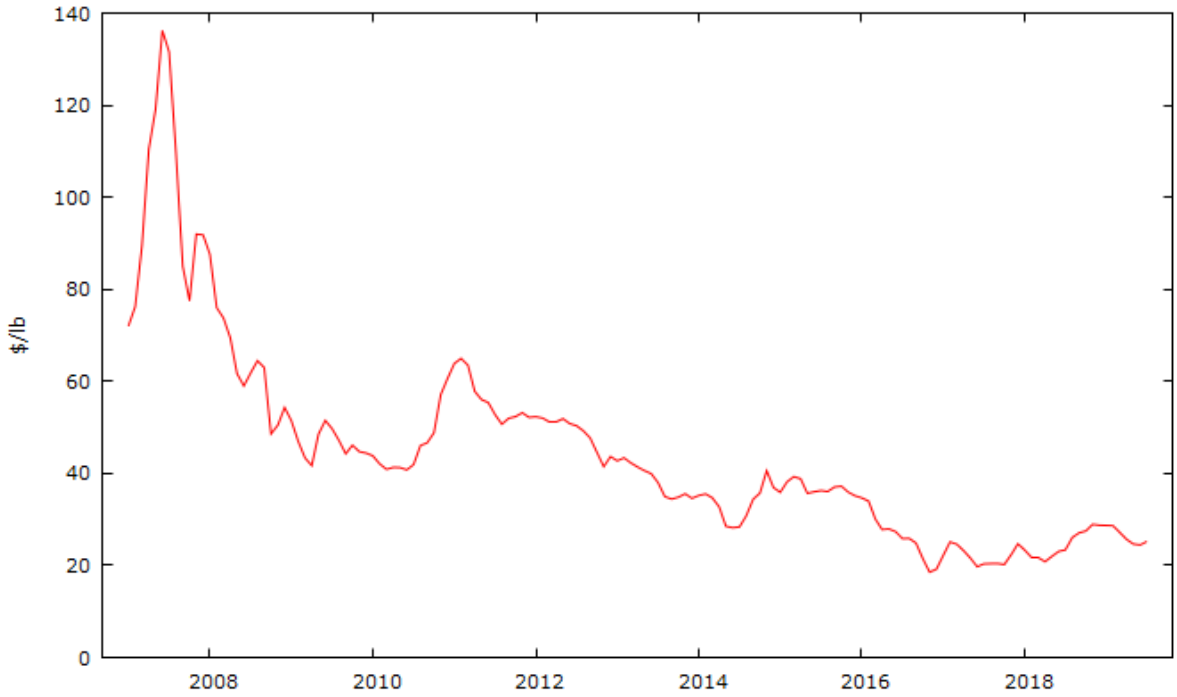



Рисунок 1.4 – Временной ряд PURAN

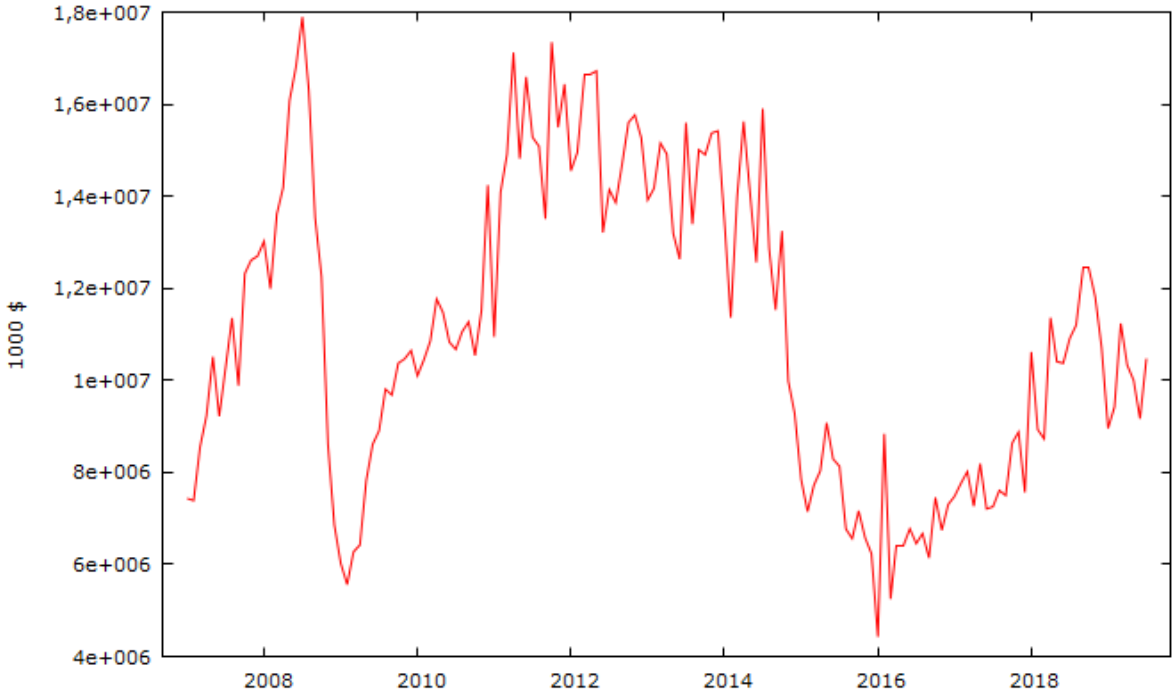



Рисунок 1.5 – Временной ряд EXOIL

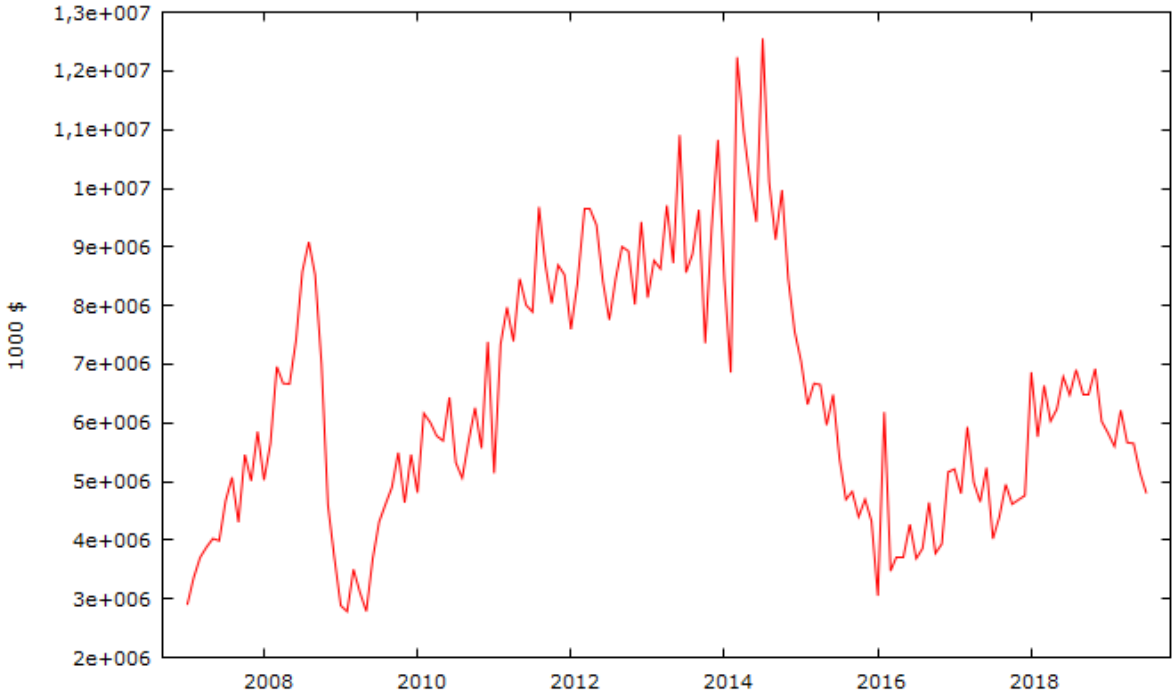



Рисунок 1.6 – Временной ряд PCRUD

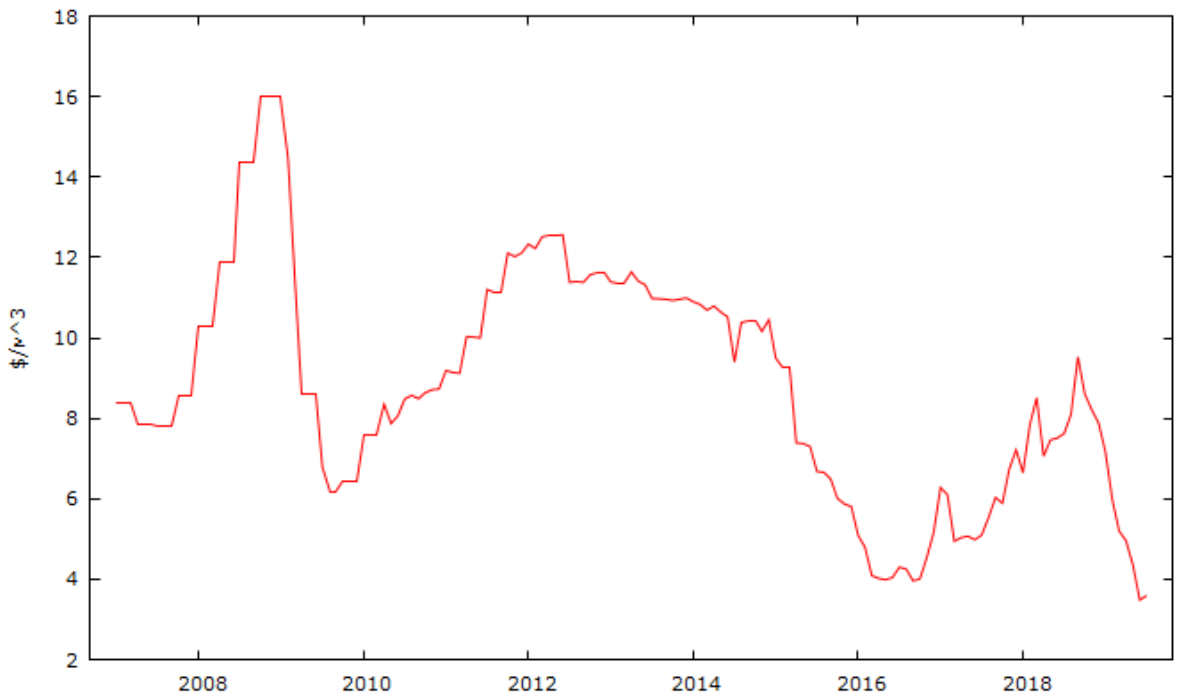



Рисунок 1.7 – Временной ряд PGAS

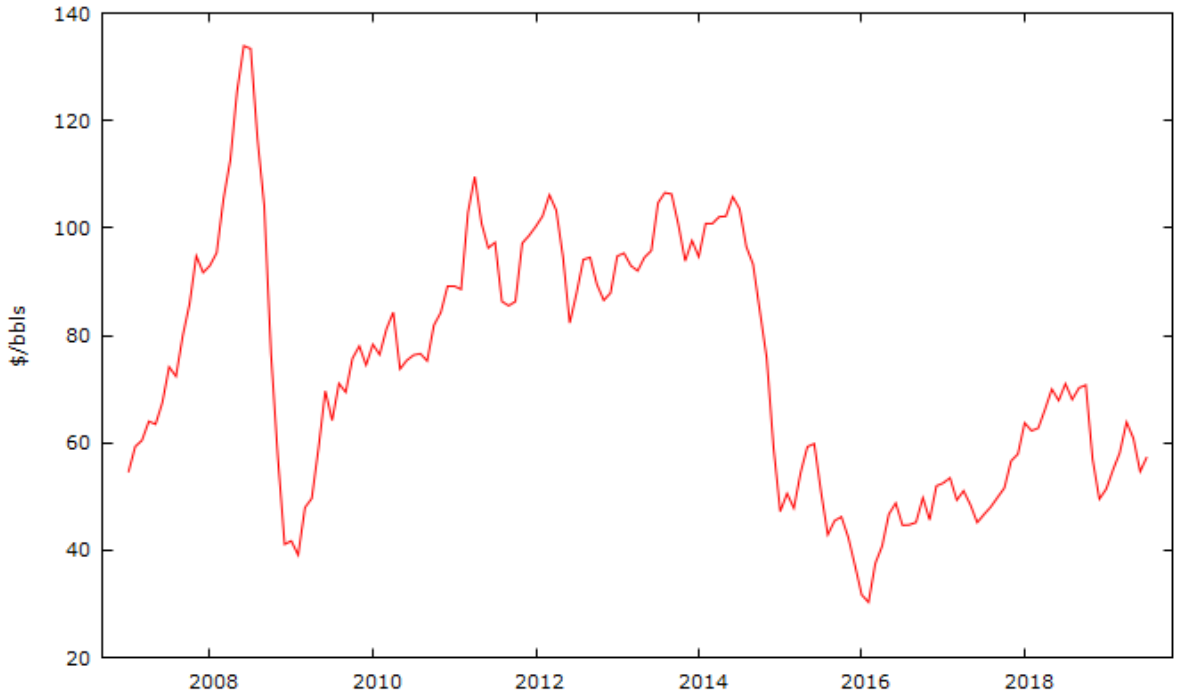



Рисунок 1.8 – Временной ряд POIL

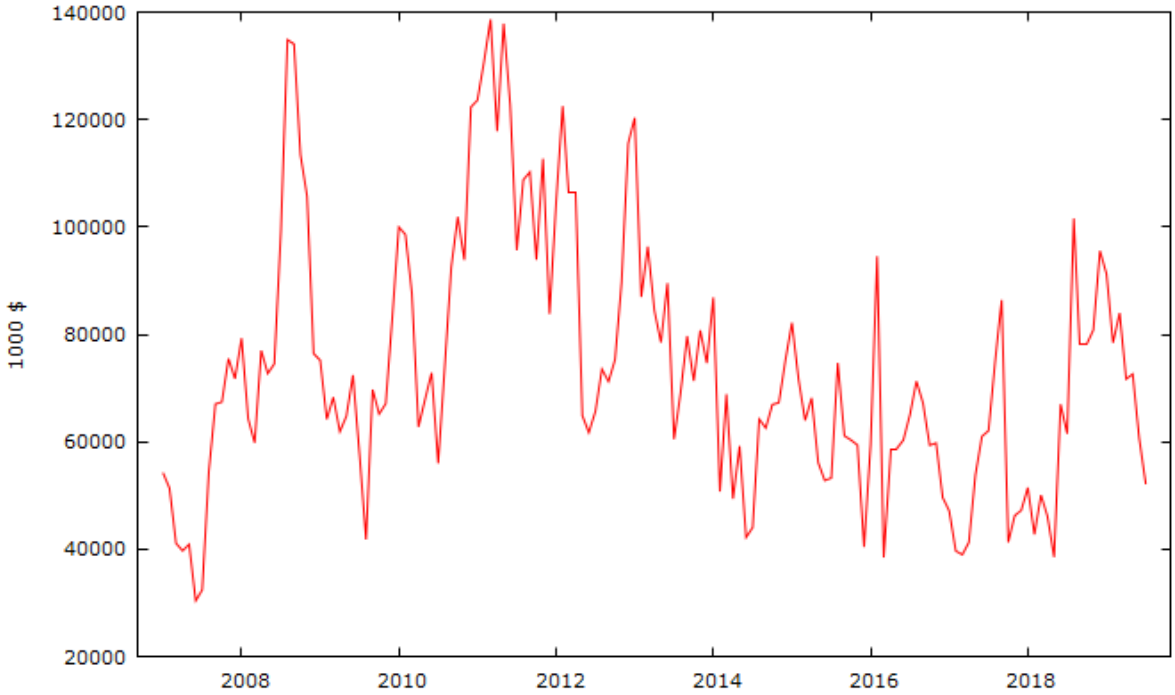



Рисунок 1.9 – Временной ряд EENER

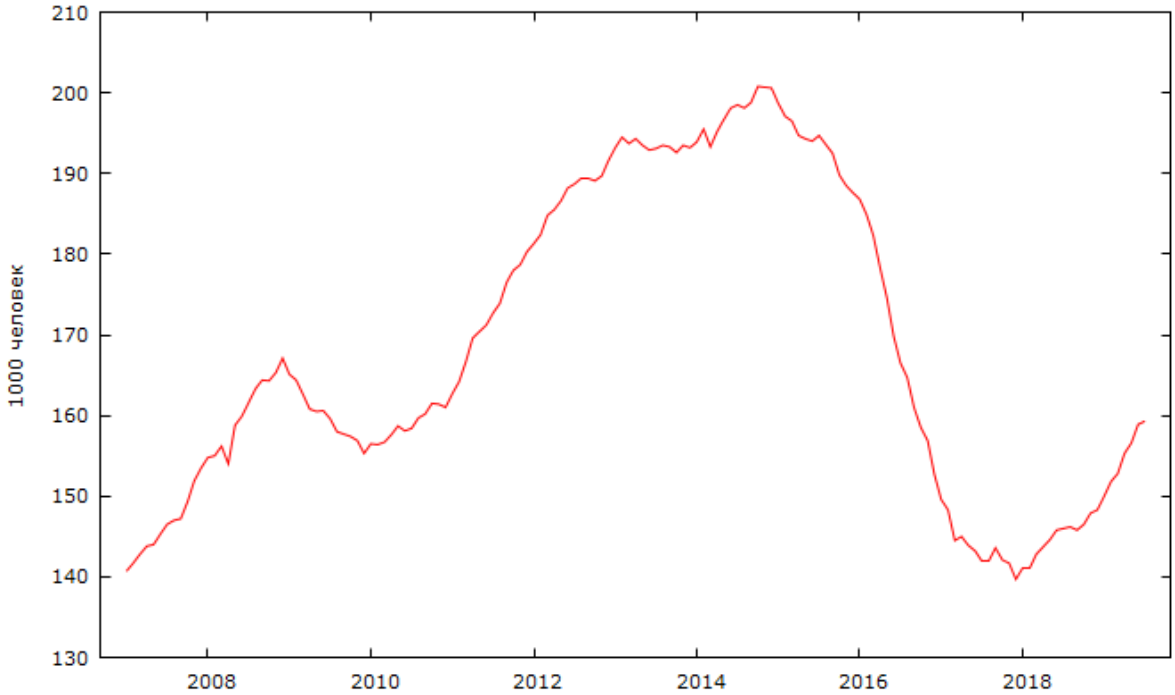



Рисунок 1.10 – Временной ряд EMP

## 1.13 Диаграммы рассеяния

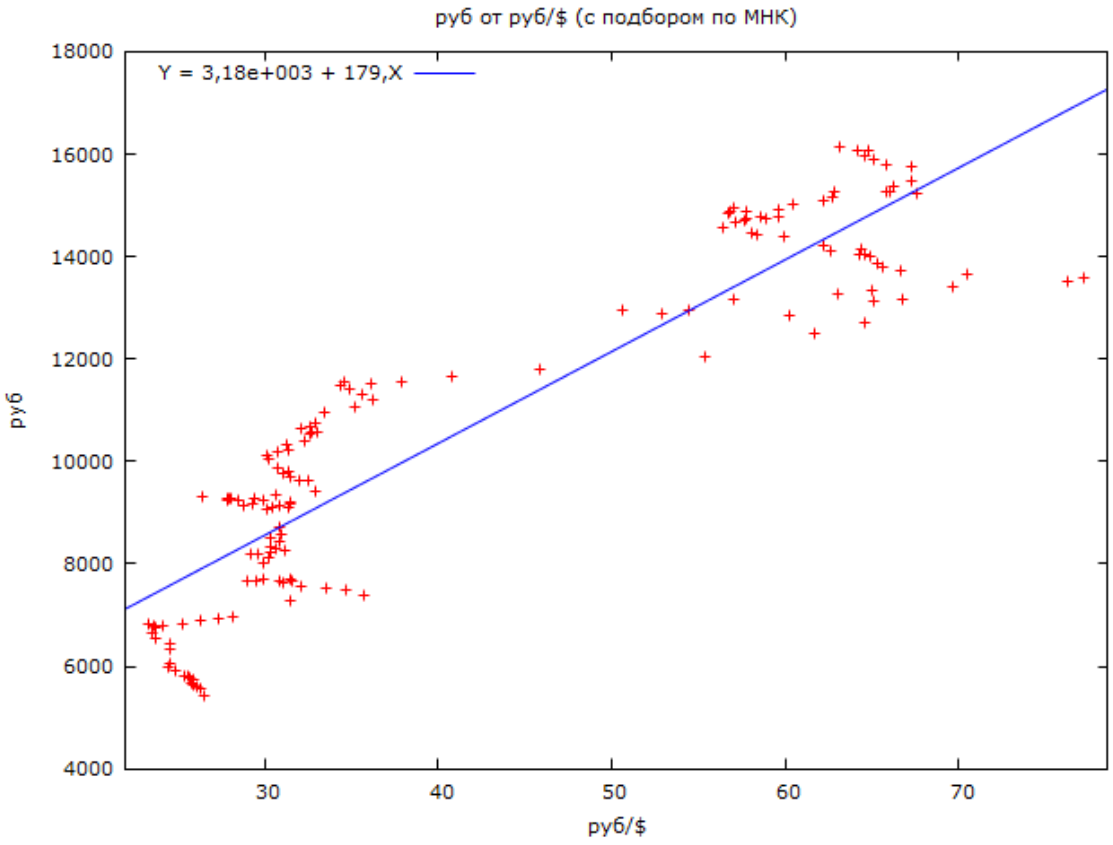



Рисунок 1.11 – Диаграмма рассеяния XXX и RU

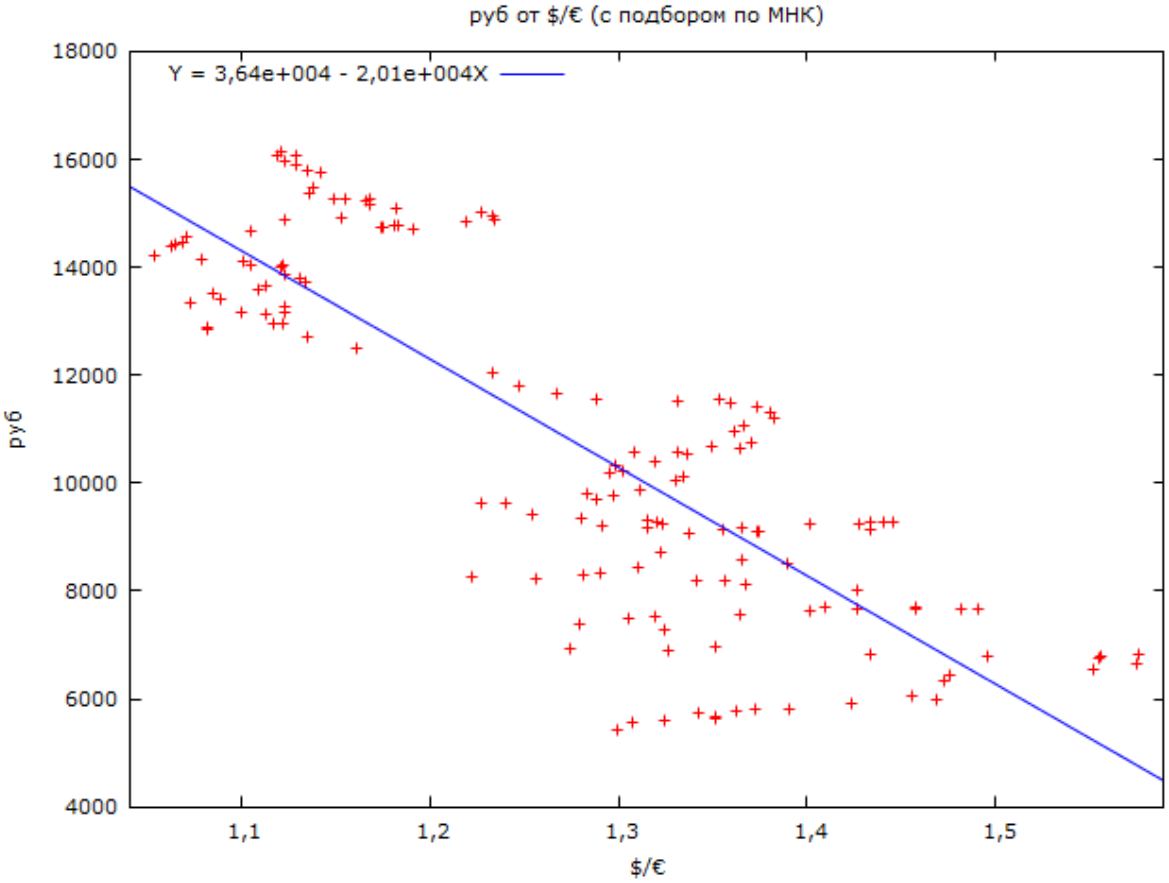



Рисунок 1.12 – Диаграмма рассеяния XXX и EUR

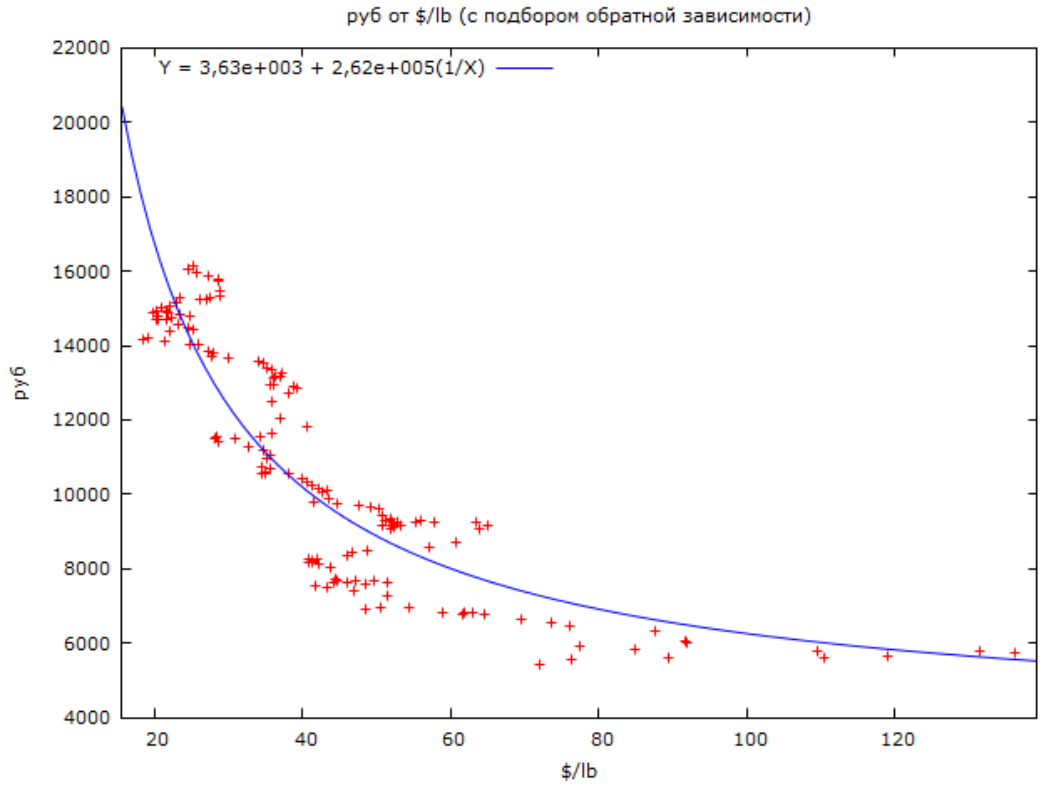



Рисунок 1.13 – Диаграмма рассеяния XXX и PURAN

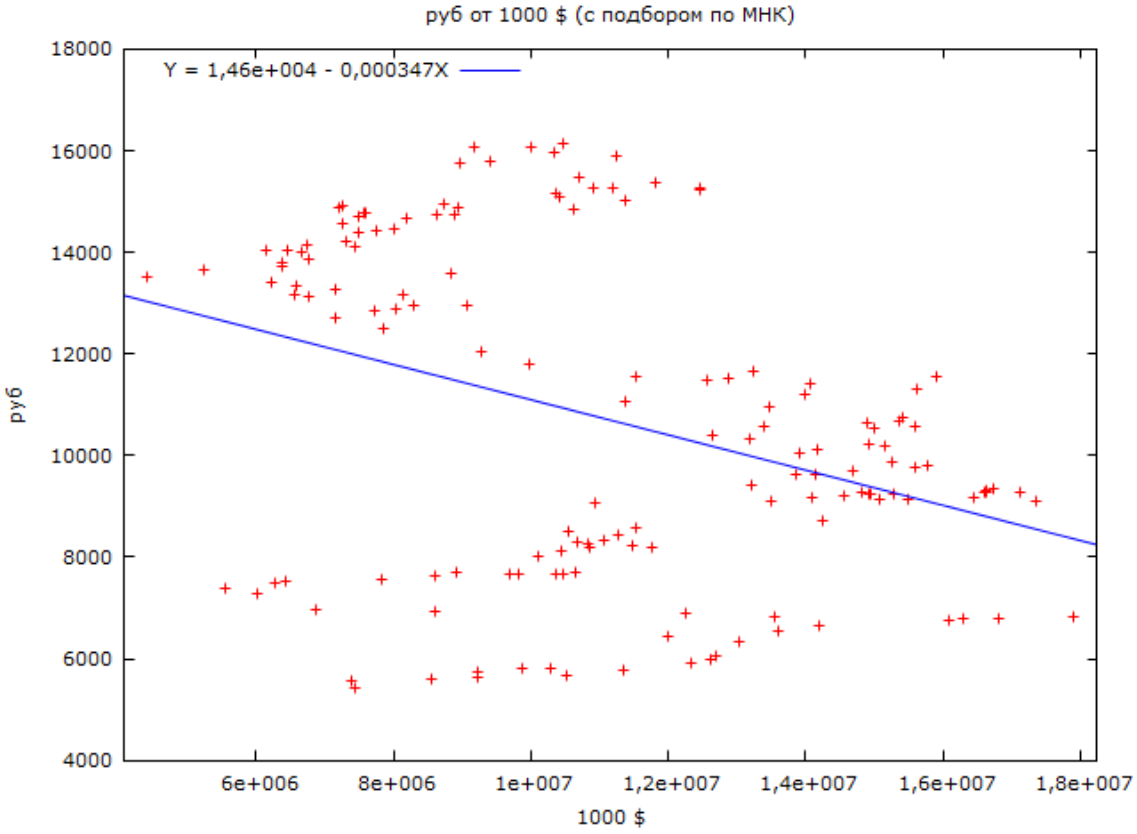



Рисунок 1.14 – Диаграмма рассеяния XXX и EXOIL

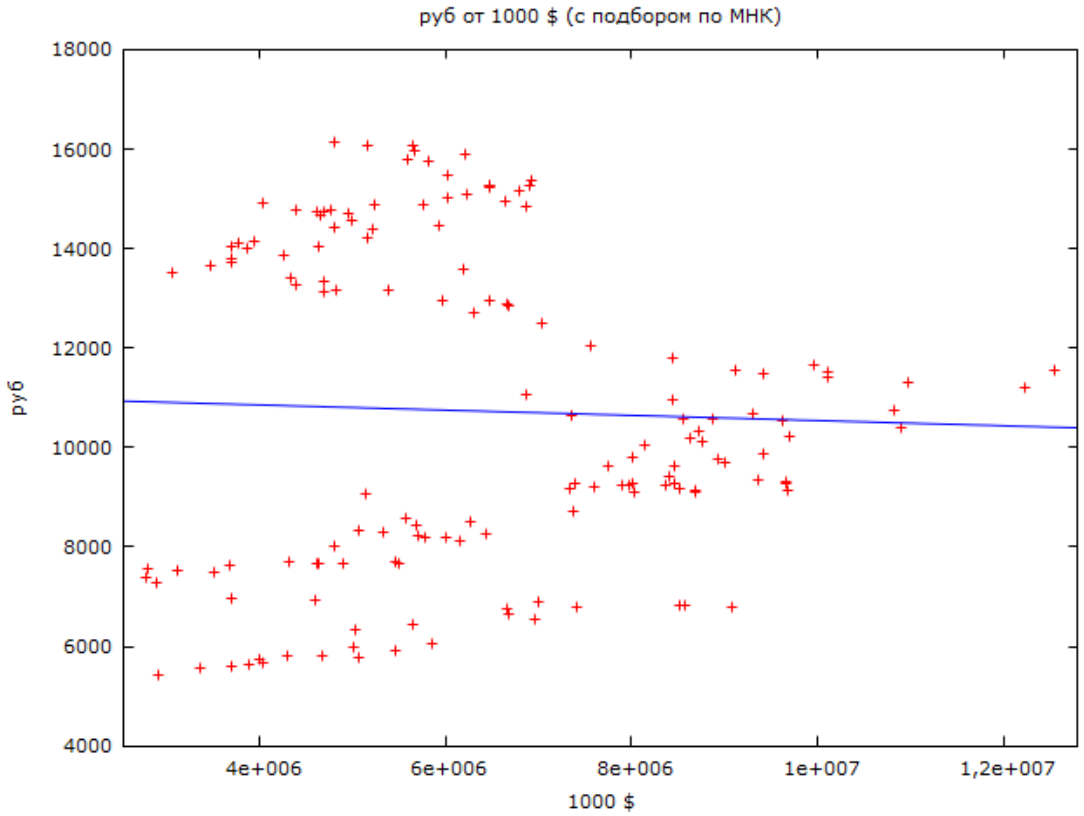



Рисунок 1.15 – Диаграмма рассеяния XXX и PCRUD

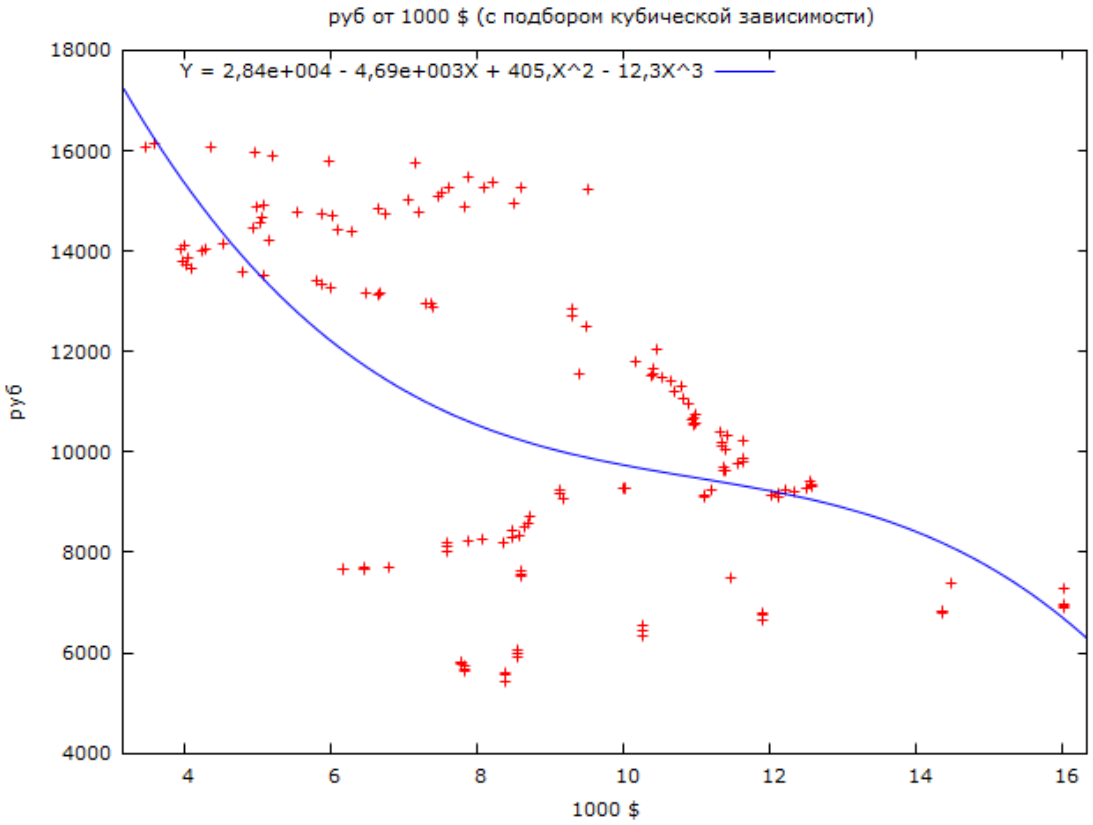



Рисунок 1.16 – Диаграмма рассеяния XXX и PGAS

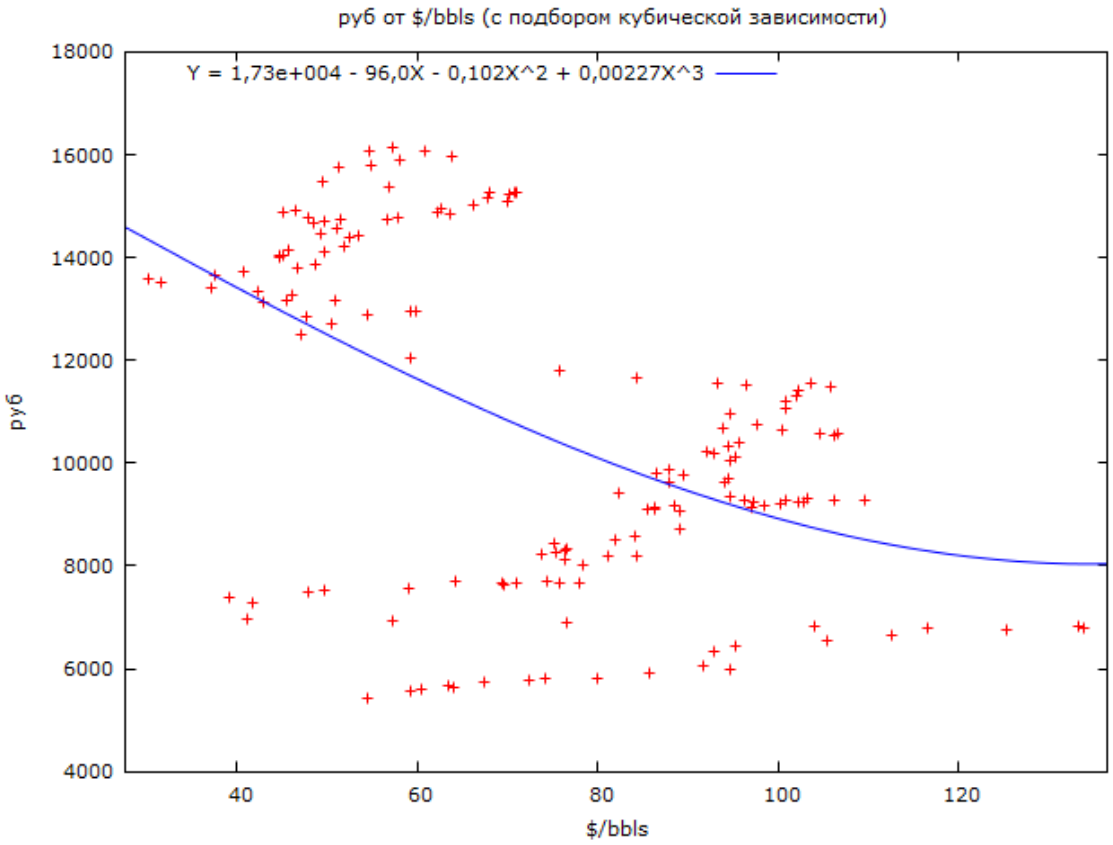



Рисунок 1.17 – Диаграмма рассеяния XXX и POIL

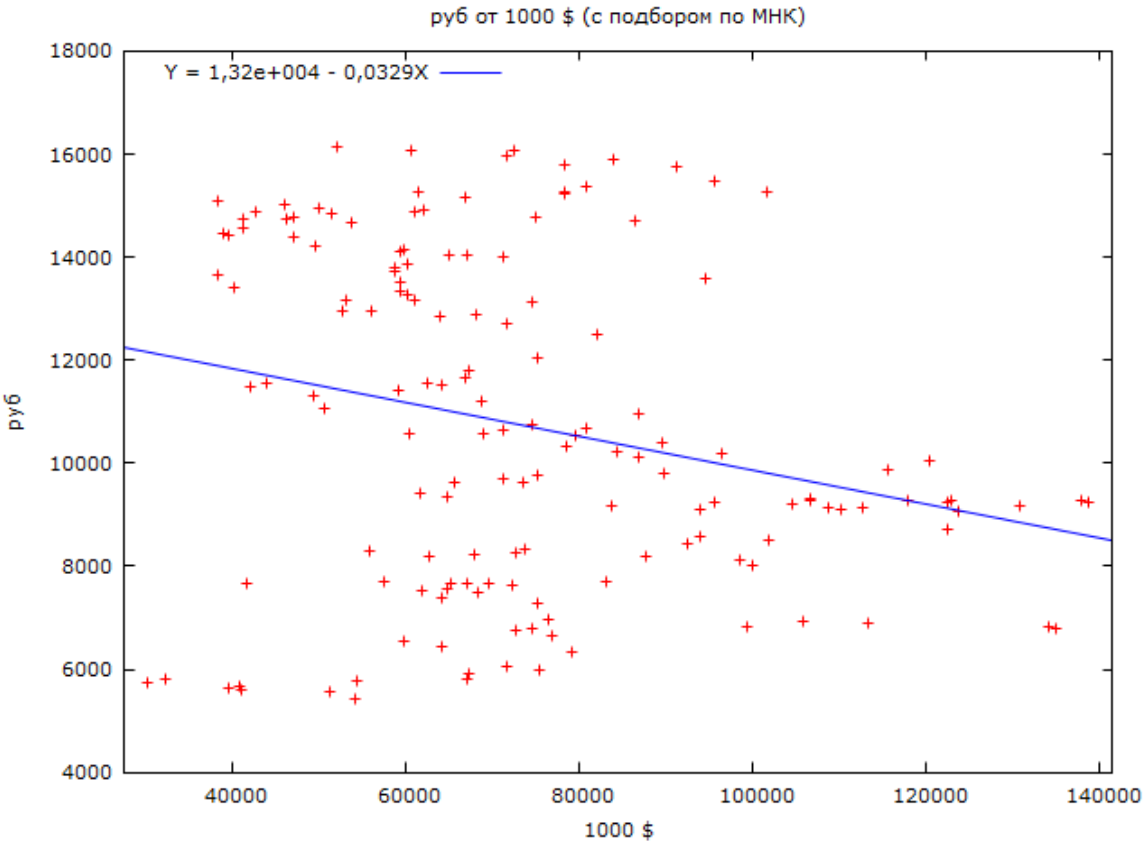



Рисунок 1.17 – Диаграмма рассеяния XXX и EENER

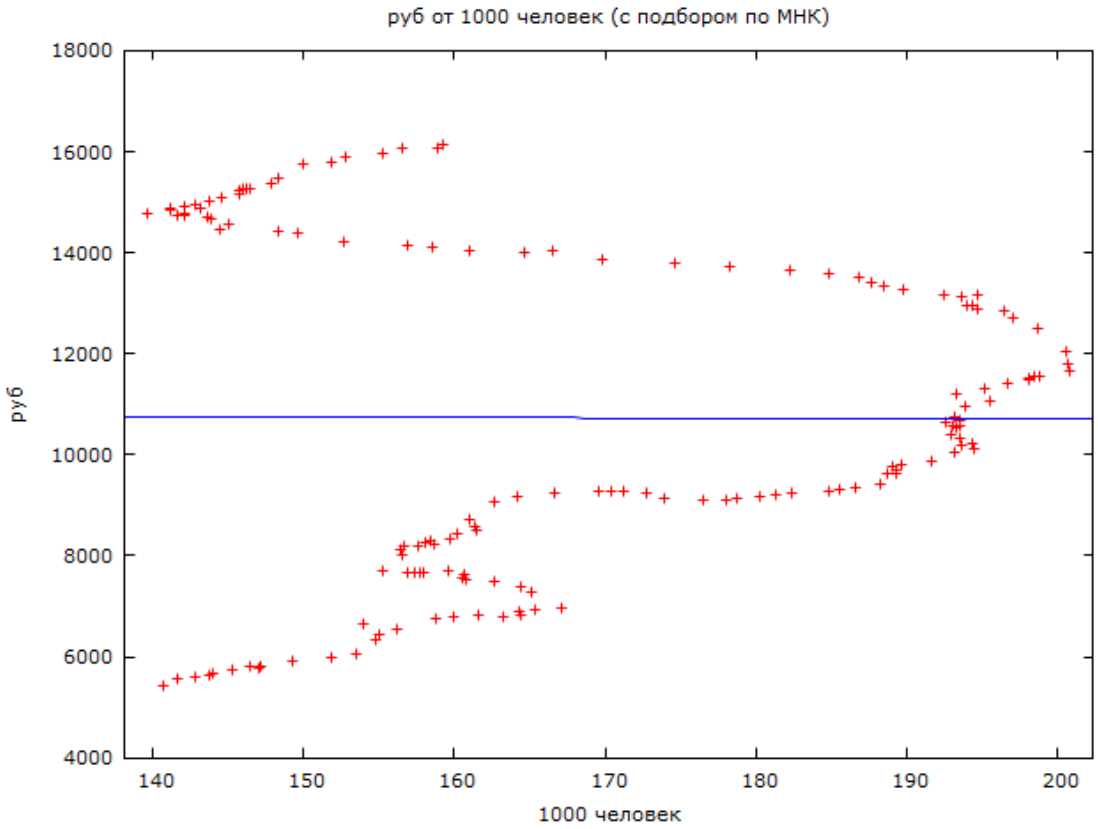



Рисунок 1.19 – Диаграмма рассеяния XXX и EMP

## 2. Построение модели и начало анализа модели

### 2.1 Выбор и построение модели множественной регрессии МНК

```
Модель 1: МНК, использованы наблюдения 2007:01-2019:07 (T = 151)
Зависимая переменная: XXX

             Коэффициент      Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const     13508,3          1695,86            7,965     4,99e-013  ***
  RU          136,384           8,74494        15,60      1,70e-032  ***
  EUR       -5330,77         1037,92           -5,136     9,16e-07   ***
  PURAN       -36,9208          3,77113        -9,790     1,37e-017  ***
  EXOIL         8,07909e-05     5,28742e-05     1,528     0,1288
  PCRUD         0,000296033     6,42394e-05     4,608     9,00e-06   ***
  PGAS        -73,9112         30,3057         -2,439     0,0160     **
  POIL         17,7512          8,06454         2,201     0,0294     **
  EENER        -0,00403776      0,00274462     -1,471     0,1435
  EMP         -19,2015          3,83478        -5,007     1,62e-06   ***

Среднее зав. перемен    10727,61   Ст. откл. зав. перемен  3196,568
Сумма кв. остатков      59315493   Ст. ошибка модели       648,5964
R-квадрат               0,961300   Испр. R-квадрат         0,958830
F(9, 141)               389,1587   P-значение (F)          6,68e-95
Лог. правдоподобие     -1186,783   Крит. Акаике            2393,566
Крит. Шварца            2423,739   Крит. Хеннана-Куинна    2405,823
Параметр rho            0,677412   Стат. Дарбина-Вотсона   0,636291

Исключая константу, наибольшее p-значение получено для переменной 9 (EENER)

Тест Рамсея (RESET) (только кубы) -
  Нулевая гипотеза: спецификация адекватна
  Тестовая статистика: F(1, 140) = 0,338925
  p-значение = P(F(1, 140) > 0,338925) = 0,561387
```

Рисунок 2.1 – Первоначальная модель МНК

### 2.2 Проверка значимости

#### 2.2.1 Тест Фишера

Нулевая гипотеза – модель в целом не значима

Таблица 2.1 – тест Фишера

| | |
|---|---|
| t-стат | 389,1587 |
| p-знач | $6{,}68 \cdot 10^{-95}$ |
| $t_{кр}$ 1% | 2,5356 |
| решение | Значима на 1%-ом уровне значимости |

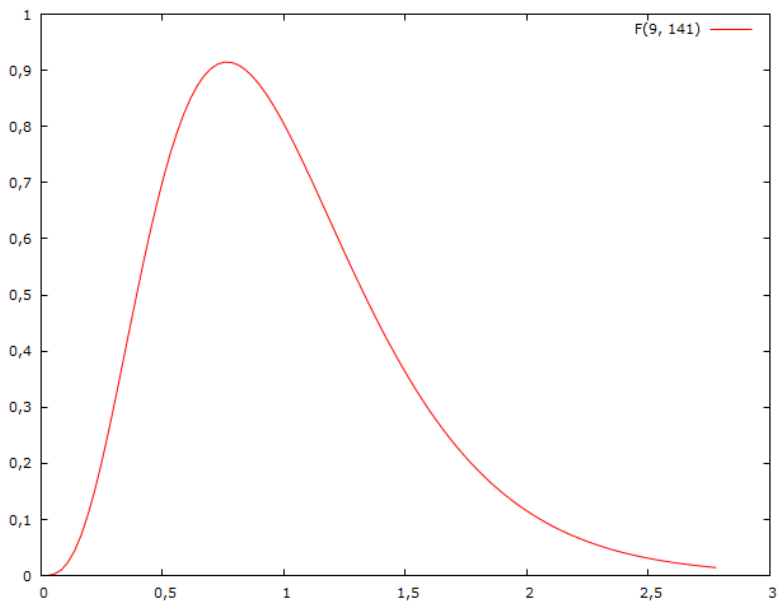



Рисунок 2.2 – распределение вероятности тестовой статистики в случае, когда модель незначима (9,141)

Модель значима  на 1%-ом уровне значимости, нулевая гипотеза отвергается

### 2.2.2 Тест Стьюдента

Нулевая гипотеза – параметр незначим.

$t_{кр}$ 10% = 1,65573, $t_{кр}$ 5% = 1,97693, $t_{кр}$ 1% = 2,61115

Таблица 2.2 – проверка значимости коэффициентов

| Регрессор | t-стат | p-знач | $t_{кр}$ | решение |
|---|---|---|---|---|
| RU | 15,60 | $1{,}70 \cdot 10^{-32}$ | $t_{кр}$ 1% = 2,61115 | Значим на 1% |
| EUR | -5,136 | $9{,}16 \cdot 10^{-7}$ | $t_{кр}$ 1% = 2,61115 | Значим на 1% |
| PURAN | -9,790 | $1{,}37 \cdot 10^{-17}$ | $t_{кр}$ 1% = 2,61115 | Значим на 1% |
| EXOIL | 1,528 | 0,128 | $t_{кр}$ 10% = 1,65573 | Не значим |
| PCRUD | 4,608 | $9{,}00 \cdot 10^{-6}$ | $t_{кр}$ 1% = 2,61115 | Значим на 1% |
| PGAS | -2,439 | 0,0160 | $t_{кр}$ 5% = 1,97693 | Значим на 5% |
| POIL | 2,201 | 0,0294 | $t_{кр}$ 5% = 1,97693 | Значим на 5% |
| EENER | -1,471 | 0,1435 | $t_{кр}$ 10% = 1,65573 | Не значим |
| EMP | -5,007 | $1{,}62 \cdot 10^{-6}$ | $t_{кр}$ 1% = 2,61115 | Значим на 1% |

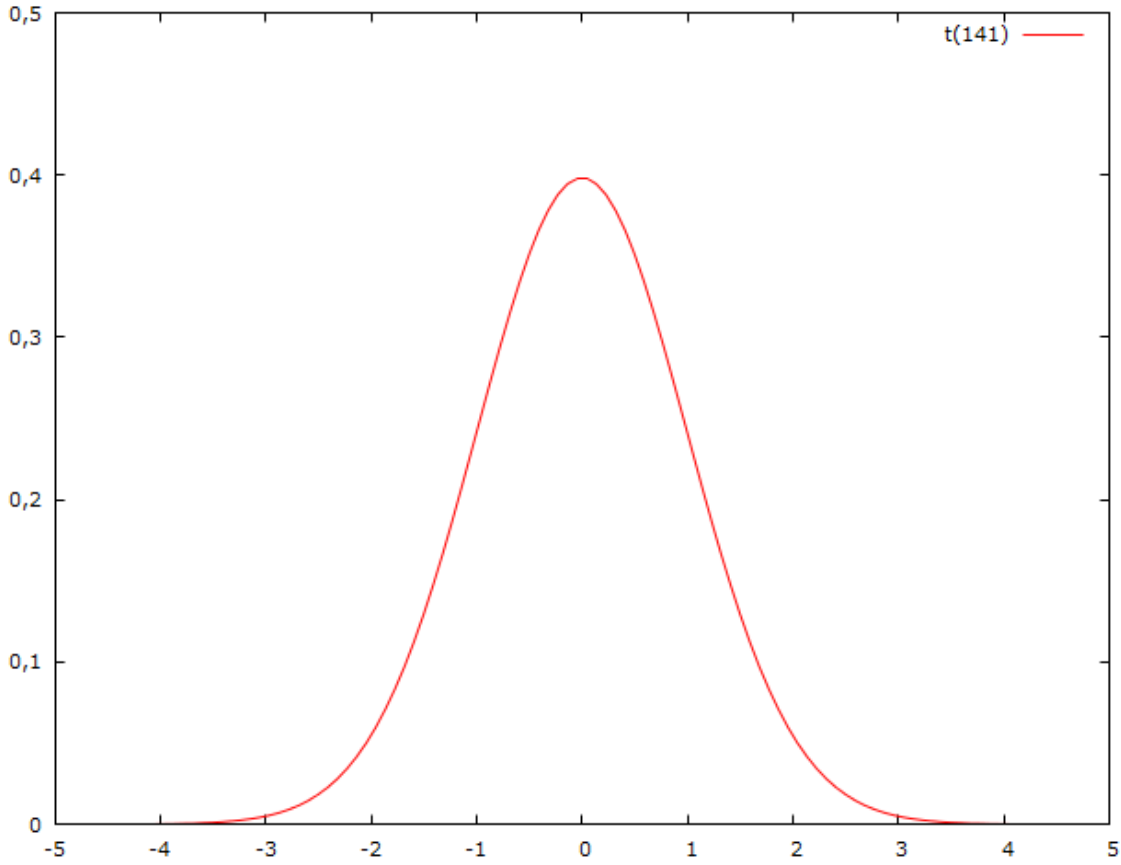



Рисунок 2.3 – распределение вероятности тестовой статистики в случае, когда коэффициент незначим (141 степень свободы)

RU, EUR, PURAN, PCRUD, EMP значимы 1-ом уровне значимости.

PGAS, POIL значимы на 5%-ом уровне значимости.

EXOIL и EENER не значимы.

## 2.3 RESET-тест

```
Вспомогательная регрессия для теста Рамсея
МНК, использованы наблюдения 2007:01-2019:07 (T = 151)
Зависимая переменная: XXX

             Коэффициент       Ст. ошибка     t-статистика  P-значение
  ---------------------------------------------------------------------
  const     12914,3           1982,60            6,514      1,22e-09   ***
  RU          117,932           32,8858          3,586      0,0005     ***
  EUR       -4920,56          1256,52           -3,916      0,0001     ***
  PURAN       -35,4977           4,50150         -7,886      8,02e-013  ***
  EXOIL         6,90591e-05      5,67005e-05      1,218      0,2253
  PCRUD         0,000267464      8,09588e-05      3,304      0,0012     ***
  PGAS        -73,7640          30,3781          -2,428      0,0164     **
  POIL         16,4709           8,37731          1,966      0,0513     *
  EENER        -0,00332016       0,00301459      -1,101      0,2726
  EMP         -15,0185           8,14876         -1,843      0,0674     *
  yhat^3        2,79396e-010     4,79919e-010     0,5822     0,5614

Тестовая статистика: F = 0,338925,
p-значение = P(F(1,140) > 0,338925) = 0,561
```

Рисунок 2.5а – тест Рамсея (только кубы)

```
Модель 1: МНК, использованы наблюдения 2007:01-2019:07 (T = 151)
Зависимая переменная: XXX

             Коэффициент       Ст. ошибка     t-статистика  P-значение
  ---------------------------------------------------------------------
  const     13508,3           1695,86            7,965      4,99e-013  ***
  RU          136,384           8,74494         15,60       1,70e-032  ***
  EUR       -5330,77          1037,92           -5,136      9,16e-07   ***
  PURAN       -36,9208          3,77113         -9,790      1,37e-017  ***
  EXOIL         8,07909e-05     5,28742e-05      1,528      0,1288
  PCRUD         0,000296033     6,42394e-05      4,608      9,00e-06   ***
  PGAS        -73,9112         30,3057          -2,439      0,0160     **
  POIL         17,7512           8,06454          2,201      0,0294     **
  EENER        -0,00403776       0,00274462      -1,471      0,1435
  EMP         -19,2015           3,83478         -5,007      1,62e-06   ***

  Среднее зав. перемен    10727,61   Ст. откл. зав. перемен  3196,568
  Сумма кв. остатков      59315493   Ст. ошибка модели       648,5964
  R-квадрат               0,961300   Испр. R-квадрат         0,958830
  F(9, 141)               389,1587   P-значение (F)          6,68e-95
  Лог. правдоподобие     -1186,783   Крит. Акаике            2393,566
  Крит. Шварца            2423,739   Крит. Хеннана-Куинна    2405,823
  Параметр rho            0,677412   Стат. Дарбина-Вотсона   0,636291

  Исключая константу, наибольшее p-значение получено для переменной 9 (EENER)

  Тест Рамсея (RESET) (только кубы) -
    Нулевая гипотеза: спецификация адекватна
    Тестовая статистика: F(1, 140) = 0,338925
    p-значение = P(F(1, 140) > 0,338925) = 0,561387
```

Рисунок 2.5б – модель МНК

Нулевая гипотеза – спецификация модели адекватна (только кубы). $t_{кр}$(10%) = 2,74171. t-стат = 0,338925. P-значение = 0,561. Значит нулевая гипотеза не отвергается, спецификация модели адекватна на 10%-ом уровне значимости.

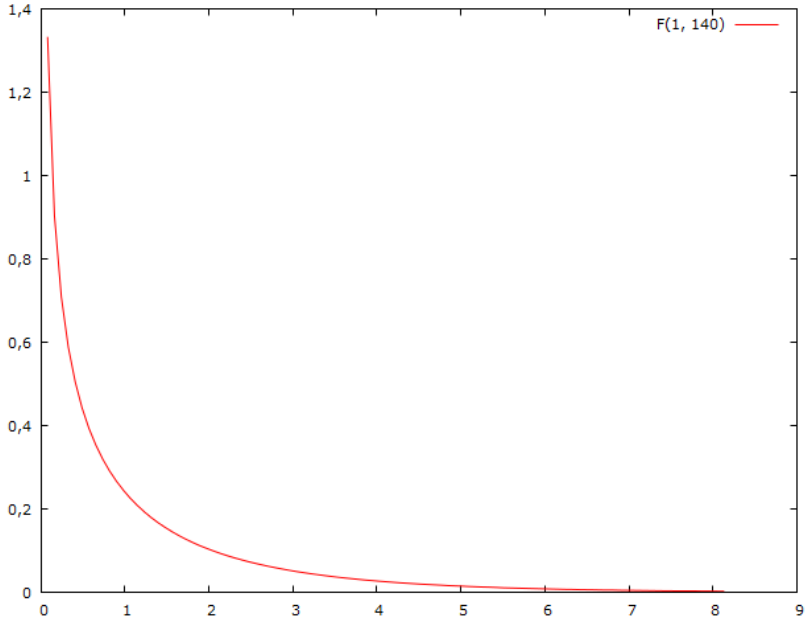



Рисунок 2.6 – распределение вероятности тестовой статистики в случае, когда спецификация модели адекватна

## 2.4 Анализ мультиколлинеарности параметров с помощью факторов инфляции дисперсии

Строится новая модель, исключающая исследуемый параметр XXX.

```
Метод инфляционных факторов
Минимальное возможное значение = 1.0
Значения > 10.0 могут указывать на наличие мультиколлинеарности

          RU    7,235
         EUR    6,551
       PURAN    2,383
       EXOIL   10,858
       PCRUD    6,670
        PGAS    2,705
        POIL   12,192
       EENER    1,543
         EMP    1,970

VIF(j) = 1/(1 - R(j)^2), где R(j) - это коэффициент множественной корреляции
между переменной j и другими независимыми переменными

Belsley-Kuh-Welsch collinearity diagnostics:

  variance proportions
```

Рисунок 2.7 – метод факторов инфляции дисперсии

Фактор инфляции дисперсии для параметров EXOIL и POIL > 10.

```
Анализ главных компонент
n = 151

Собственные значения для матрицы корреляций

Компонента  Собс. знач.  Доля   Интегральная
    1       1,9174       0,9587       0,9587
    2       0,0826       0,0413       1,0000

Собственные векторы (нагрузка на компоненты)

                PC1      PC2
EXOIL         0,707   -0,707
POIL          0,707    0,707
```

Рисунок 2.8 – анализ главных параметров

Строим новую модель, где регрессоры EXOIL и POIL заменены на собственные векторы PC1 и PC2.

```
Модель 3: МНК, использованы наблюдения 2007:01-2019:07 (T = 151)
Зависимая переменная: XXX

              Коэффициент      Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const     15720,0          1801,36            8,727      6,70e-015  ***
  RU          136,384           8,74494        15,60       1,70e-032  ***
  EUR       -5330,77         1037,92           -5,136      9,16e-07   ***
  PURAN       -36,9208          3,77113        -9,790      1,37e-017  ***
  PCRUD         0,000296033     6,42394e-05     4,608      9,00e-06   ***
  PGAS        -73,9112         30,3057         -2,439      0,0160     **
  EENER        -0,00403776      0,00274462     -1,471      0,1435
  EMP         -19,2015          3,83478        -5,007      1,62e-06   ***
  PC1         476,354         111,015           4,291      3,29e-05   ***
  PC2          99,2638        228,739           0,4340     0,6650

Среднее зав. перемен    10727,61   Ст. откл. зав. перемен  3196,568
Сумма кв. остатков      59315493   Ст. ошибка модели       648,5964
R-квадрат               0,961300   Испр. R-квадрат         0,958830
F(9, 141)               389,1587   P-значение (F)          6,68e-95
Лог. правдоподобие     -1186,783   Крит. Акаике            2393,566
Крит. Шварца            2423,739   Крит. Хеннана-Куинна    2405,823
Параметр rho            0,677412   Стат. Дарбина-Вотсона   0,636291

Исключая константу, наибольшее p-значение получено для переменной 15 (PC2)
```

Рисунок 2.9 – модель МНК с собственными векторами PC1 и PC2 (EXOIL и POIL заменены)

```
Метод инфляционных факторов
Минимальное возможное значение = 1.0
Значения > 10.0 могут указывать на наличие мультиколлинеарности

          RU    7,235
         EUR    6,551
       PURAN    2,383
       PCRUD    6,670
        PGAS    2,705
       EENER    1,543
         EMP    1,970
         PC1    8,426
         PC2    1,542

VIF(j) = 1/(1 - R(j)^2), где R(j) - это коэффициент множественной корреляции
между переменной j и другими независимыми переменными

Belsley-Kuh-Welsch collinearity diagnostics:

  variance proportions
```

Рисунок 2.10 –факторы инфляции дисперсии для новой модели

Факторы инфляции дисперсии для всех параметров меньше 10.

## 2.4 Проверка на избыточные переменные

```
Модель 1: МНК, использованы наблюдения 2007:01-2019:07 (T = 151)
Зависимая переменная: XXX

              Коэффициент      Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const      15720,0          1801,36            8,727      6,70e-015  ***
  RU           136,384           8,74494        15,60       1,70e-032  ***
  EUR        -5330,77         1037,92           -5,136      9,16e-07   ***
  PURAN        -36,9208          3,77113        -9,790      1,37e-017  ***
  PCRUD          0,000296033     6,42394e-05     4,608      9,00e-06   ***
  PGAS         -73,9112         30,3057         -2,439      0,0160     **
  EENER         -0,00403776      0,00274462     -1,471      0,1435
  EMP          -19,2015          3,83478        -5,007      1,62e-06   ***
  PC1          476,354         111,015           4,291      3,29e-05   ***
  PC2           99,2638        228,739           0,4340     0,6650

Среднее зав. перемен     10727,61   Ст. откл. зав. перемен  3196,568
Сумма кв. остатков       59315493   Ст. ошибка модели       648,5964
R-квадрат                0,961300   Испр. R-квадрат         0,958830
F(9, 141)                389,1587   P-значение (F)          6,68e-95
Лог. правдоподобие      -1186,783   Крит. Акаике            2393,566
Крит. Шварца             2423,739   Крит. Хеннана-Куинна    2405,823
Параметр rho             0,677412   Стат. Дарбина-Вотсона   0,636291

Исключая константу, наибольшее p-значение получено для переменной 15 (PC2)
```

Рисунок 2.11а – «длинная» модель МНК

```
Модель 1   Модель 2

Модель 2: МНК, использованы наблюдения 2007:01-2019:07 (T = 151)
Зависимая переменная: XXX

              Коэффициент      Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const      15426,8          1770,01            8,716      6,58e-015  ***
  RU           134,753           8,70052        15,49       2,19e-032  ***
  EUR        -5251,83         980,423           -5,357      3,31e-07   ***
  PURAN        -36,8008          3,70397        -9,935      5,18e-018  ***
  PCRUD          0,000298946     6,35607e-05     4,703      5,97e-06   ***
  PGAS         -87,9974         29,2644         -3,007      0,0031     ***
  EMP          -18,8247          3,78448        -4,974      1,85e-06   ***
  PC1          436,931         108,934           4,011      9,71e-05   ***

Среднее зав. перемен     10727,61   Ст. откл. зав. перемен  3196,568
Сумма кв. остатков       60551004   Ст. ошибка модели       650,7178
R-квадрат                0,960494   Испр. R-квадрат         0,958560
F(7, 143)                496,6731   P-значение (F)          5,74e-97
Лог. правдоподобие      -1188,339   Крит. Акаике            2392,679
Крит. Шварца             2416,817   Крит. Хеннана-Куинна    2402,485
Параметр rho             0,655679   Стат. Дарбина-Вотсона   0,679877
```

Рисунок 2.11б – «короткая» модель МНК

```
Тестирование модели 1:

  Нулевая гипотеза: параметры регрессии нулевые
    EENER, PC2
  Тестовая статистика: F(2, 141) = 1,46848, P-значение 0,233776
  Omitting variables improved 3 of 3 information criteria.

Модель 2: МНК, использованы наблюдения 2007:01-2019:07 (T = 151)
Зависимая переменная: XXX

             Коэффициент      Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const     15426,8          1770,01            8,716     6,58e-015  ***
  RU          134,753           8,70052        15,49      2,19e-032  ***
  EUR       -5251,83          980,423         -5,357     3,31e-07   ***
  PURAN       -36,8008          3,70397        -9,935     5,18e-018  ***
  PCRUD         0,000298946     6,35607e-05     4,703     5,97e-06   ***
  PGAS        -87,9974         29,2644         -3,007     0,0031     ***
  EMP         -18,8247          3,78448        -4,974     1,85e-06   ***
  PC1         436,931         108,934           4,011     9,71e-05   ***

Среднее зав. перемен    10727,61   Ст. откл. зав. перемен  3196,568
Сумма кв. остатков      60551004   Ст. ошибка модели       650,7178
R-квадрат               0,960494   Испр. R-квадрат         0,958560
F(7, 143)               496,6731   P-значение (F)          5,74e-97
Лог. правдоподобие     -1188,339   Крит. Акаике            2392,679
Крит. Шварца            2416,817   Крит. Хеннана-Куинна    2402,485
Параметр rho            0,655679   Стат. Дарбина-Вотсона   0,679877
```

Рисунок 2.11в – тест на избыточные переменные

$H_0$ – параметры EENER и PC2 являются незначимыми. $t_{кр}(10\%)$ = 2,3406. t-стат = 1,46848. P-значение = 0,233776. Значит нулевая гипотеза не отвергается, параметры EENER и PC2 являются незначимыми на 10%-ом уровне значимости. Исключаем EENER и PC2

```
F(2, 141)
 Правосторонняя вероятность = 0,1
 Дополняющая вероятность = 0,9

 Критическое значение = 2,3406
```

Рисунок 2.12 – Расчет критического значения

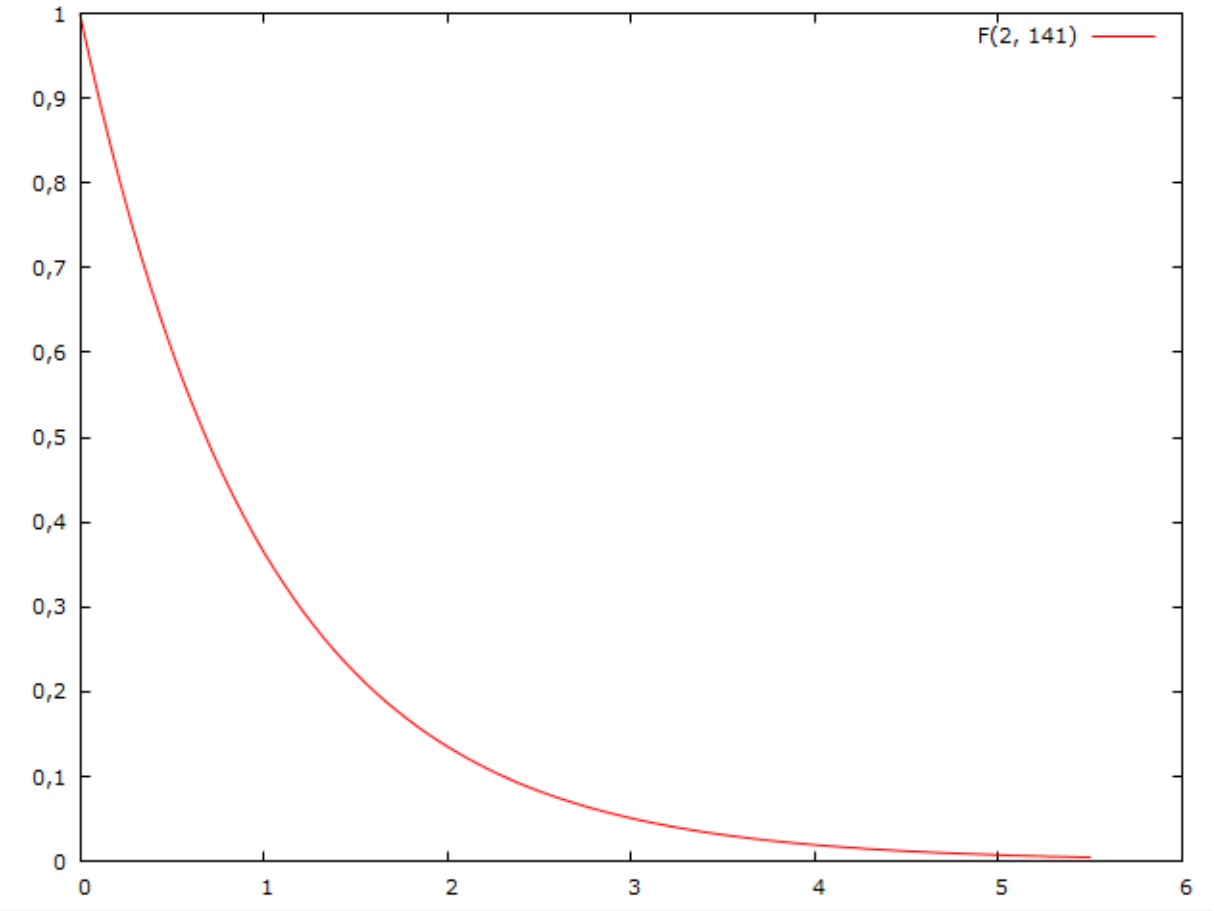



Рисунок 2.13 – Распределение F-статистики

Теперь необходимо провести RESET тест для «короткой модели»

```
Вспомогательная регрессия для теста Рамсея
МНК, использованы наблюдения 2007:01-2019:07 (T = 151)
Зависимая переменная: XXX

              Коэффициент       Ст. ошибка     t-статистика  P-значение
  ---------------------------------------------------------------------
  const      14766,0           2164,77            6,821      2,40e-010  ***
  RU           118,877           31,0437          3,829      0,0002     ***
  EUR        -4891,49          1193,04           -4,100      6,92e-05   ***
  PURAN        -35,7606           4,19509        -8,524      2,05e-014  ***
  PCRUD          0,000273466      7,96648e-05     3,433      0,0008     ***
  PGAS         -86,1146          29,5500          -2,914      0,0041     ***
  EMP          -15,2749           7,66597         -1,993      0,0482     **
  PC1          400,106          129,235            3,096      0,0024     ***
  yhat^3         2,39829e-010     4,50056e-010     0,5329     0,5949

Тестовая статистика: F = 0,283969,
p-значение = P(F(1,142) > 0,283969) = 0,595
```

Рисунок 2.14 – тест Рамсея (только кубы)

Нулевая гипотеза – спецификация модели адекватна (только кубы). $t_{кр}$(10%) = 2,74119. t-стат = 0,283969. P-значение = 0,595. Значит нулевая гипотеза не отвергается, спецификация модели адекватна на 10%-ом уровне значимости.

### 2.5 Анализ структурного сдвига

```
Тест на отношение правдоподобия Квандта (Quandt) для структурных изменений в неизвестной точке,
с 15-процентным цензурированием:
Максимум F(8, 135) = 42,9722 достигается для наблюдения 2014:12
Asymptotic p-value = 2,99309e-071 for chi-square(8) = 343,777
```

Рисунок 2.15 – тест на отношения правдоподобия Квандта

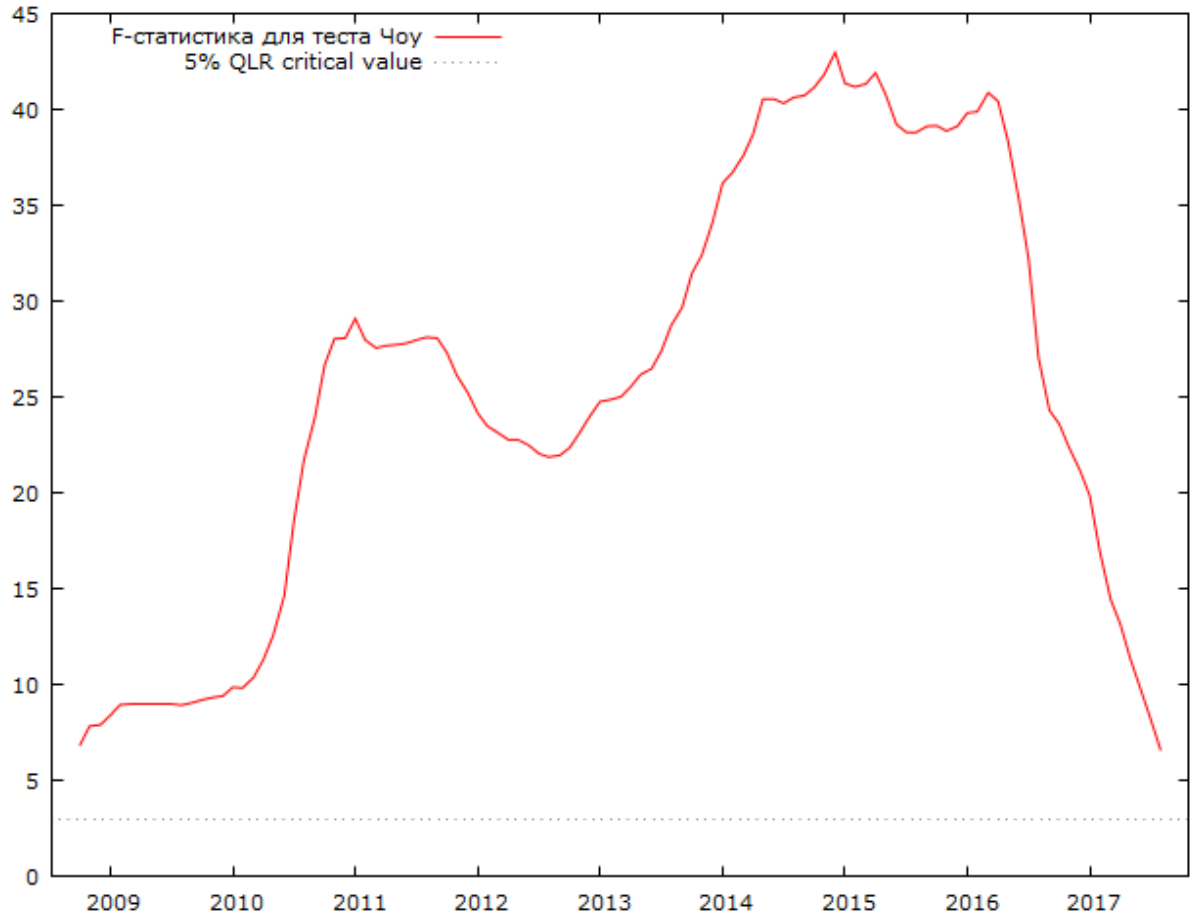



Рисунок 2.16 – F-статистика для теста Чоу

Максимум достигается для наблюдения 2014:12

```
Расширенная регрессия для теста Чоу
МНК, использованы наблюдения 2007:01-2019:07 (T = 151)
Зависимая переменная: XXX

              Коэффициент      Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const     -1580,97         1505,58            -1,050     0,2956
  RU          117,412          22,8360           5,142     9,37e-07   ***
  EUR       -1016,64          663,599           -1,532     0,1279
  PURAN       -13,3461          2,60725         -5,119     1,04e-06   ***
  PCRUD         6,20750e-05     4,38871e-05      1,414     0,1595
  PGAS       -130,272          21,2381          -6,134     8,90e-09   ***
  EMP          55,1350          6,69354          8,237     1,35e-013  ***
  PC1         254,330          79,4309           3,202     0,0017     ***
  splitdum  20946,6          2402,07             8,720     9,01e-015  ***
  sd_RU       -80,9412         25,1410          -3,219     0,0016     ***
  sd_EUR     1814,46          1547,77             1,172     0,2431
  sd_PURAN     84,7859         24,6327           3,442     0,0008     ***
  sd_PCRUD     -3,95984e-05     0,000101824     -0,3889    0,6980
  sd_PGAS    -138,860          56,2220          -2,470     0,0148     **
  sd_EMP     -102,470           9,64911        -10,62      1,59e-019  ***
  sd_PC1      467,942         164,059            2,852     0,0050     ***

Среднее зав. перемен    10727,61   Ст. откл. зав. перемен  3196,568
Сумма кв. остатков      17073457   Ст. ошибка модели       355,6263
R-квадрат               0,988861   Испр. R-квадрат         0,987623
F(15, 135)              798,9423   P-значение (F)          8,0e-124
Лог. правдоподобие     -1092,759   Крит. Акаике            2217,519
Крит. Шварца            2265,795   Крит. Хеннана-Куинна    2237,131
Параметр rho            0,689471   Стат. Дарбина-Вотсона   0,623772

Тест Чоу для структурных изменений в точке 2014:12
  F(8, 135) = 42,9722 p-значение 0,0000
```

Рисунок 2.17 –теста Чоу для структурных изменений в точке 2014:12

Нулевая гипотеза – структурного сдвига нет. $t_{кр}(1\%) = 2{,}64564$. t-стат = 42,9722. P - значение = 0. Значит нулевая гипотеза отвергается, структурный сдвиг есть на 1%-ом уровне значимости.

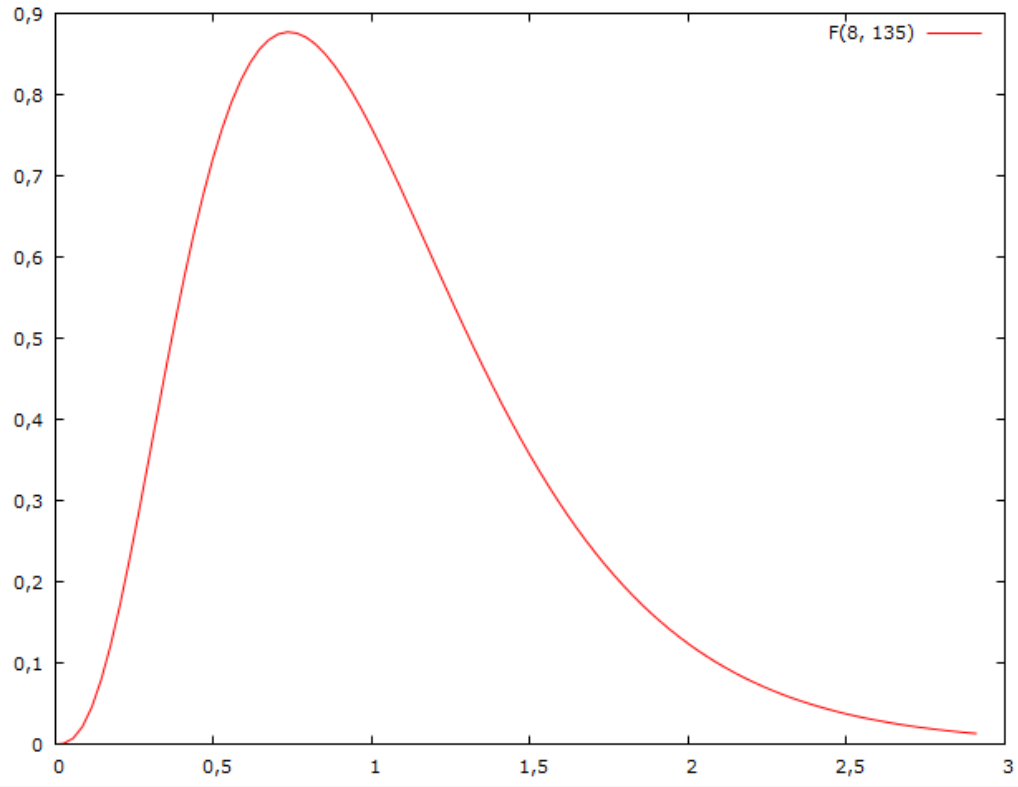



Рисунок 2.18 – распределение вероятности в случае отсутствия структурного сдвига

## 2.5 Построение и анализ модели после структурного сдвига

### 2.5.1 Выбор и построение модели множественной регрессии МНК

```
Модель 2: МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: XXX

              Коэффициент      Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const      16333,3          1816,05          8,994       1,06e-011  ***
  RU            31,1444         12,3630        2,519       0,0153     **
  EUR          612,657        1358,34          0,4510      0,6541
  PURAN         70,4400         24,7226        2,849       0,0065     ***
  EXOIL          0,000135275     6,88526e-05   1,965       0,0555     *
  PCRUD         -2,58548e-06     9,68836e-05  -0,02669     0,9788
  PGAS        -270,711          53,2875       -5,080       6,71e-06   ***
  POIL          26,3846         11,5867        2,277       0,0275     **
  EENER          0,00501844      0,00386844    1,297       0,2010
  EMP          -47,2522          7,00216      -6,748       2,18e-08   ***

Среднее зав. перемен    14365,57   Ст. откл. зав. перемен  1025,097
Сумма кв. остатков       5422632   Ст. ошибка модели       343,3414
R-квадрат               0,906175   Испр. R-квадрат         0,887818
F(9, 46)                49,36391   P-значение (F)          1,15e-20
Лог. правдоподобие     -400,9213   Крит. Акаике            821,8426
Крит. Шварца            842,0961   Крит. Хеннана-Куинна    829,6948
Параметр rho            0,472851   Стат. Дарбина-Вотсона   1,038005

Исключая константу, наибольшее p-значение получено для переменной 6 (PCRUD)
```

Рисунок 2.19 – модель МНК после анализа структурного сдвига

### 2.5.2 RESET-тест

```
Вспомогательная регрессия для теста Рамсея
МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: XXX

              Коэффициент        Ст. ошибка     t-статистика  P-значение
  ---------------------------------------------------------------------
  const     150112             512728               0,2928       0,7711
  RU           340,707           1361,02             0,2503       0,8035
  EUR         6938,53           27165,5              0,2554       0,7996
  PURAN        771,560           3074,10             0,2510       0,8030
  EXOIL          0,00144774        0,00591203        0,2449       0,8077
  PCRUD         -4,95554e-05       0,000126844      -0,3907       0,6979
  PGAS       -2993,77           11839,1             -0,2529       0,8015
  POIL         284,767           1150,36             0,2475       0,8056
  EENER          0,0532417         0,219583          0,2425       0,8095
  EMP         -525,162           2068,72             -0,2539       0,8008
  yhat^2        -0,000941637       0,00314353       -0,2995       0,7659
  yhat^3         2,76008e-08       7,50306e-08       0,3679       0,7147

Тестовая статистика: F = 3,279104,
p-значение = P(F(2,44) > 3,2791) = 0,047
```

Рисунок 2.20а – тест Рамсея (квадраты и кубы)

```
Модель 2: МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: XXX

              Коэффициент      Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------
  const     16333,3          1816,05            8,994      1,06e-011  ***
  RU           31,1444         12,3630          2,519      0,0153     **
  EUR         612,657        1358,34            0,4510     0,6541
  PURAN        70,4400         24,7226          2,849      0,0065     ***
  EXOIL         0,000135275     6,88526e-05     1,965      0,0555     *
  PCRUD        -2,58548e-06     9,68836e-05    -0,02669    0,9788
  PGAS       -270,711          53,2875         -5,080      6,71e-06   ***
  POIL         26,3846         11,5867          2,277      0,0275     **
  EENER         0,00501844      0,00386844      1,297      0,2010
  EMP         -47,2522          7,00216        -6,748      2,18e-08   ***

Среднее зав. перемен    14365,57   Ст. откл. зав. перемен  1025,097
Сумма кв. остатков       5422632   Ст. ошибка модели       343,3414
R-квадрат               0,906175   Испр. R-квадрат         0,887818
F(9, 46)                49,36391   P-значение (F)          1,15e-20
Лог. правдоподобие     -400,9213   Крит. Акаике            821,8426
Крит. Шварца            842,0961   Крит. Хеннана-Куинна    829,6948
Параметр rho            0,472851   Стат. Дарбина-Вотсона   1,038005

Исключая константу, наибольшее p-значение получено для переменной 6 (PCRUD)

Тест Рамсея (RESET) -
  Нулевая гипотеза: спецификация адекватна
  Тестовая статистика: F(2, 44) = 3,2791
  p-значение = P(F(2, 44) > 3,2791) = 0,0470481
```

Рисунок 2.20б – модель МНК

Нулевая гипотеза – спецификация модели адекватна (квадраты и кубы). $t_{кр}(1\%)$ = 5,12263. t-стат = 3,2791. Значит нулевая гипотеза не отвергается, спецификация модели адекватна на 1%-ом уровне значимости.

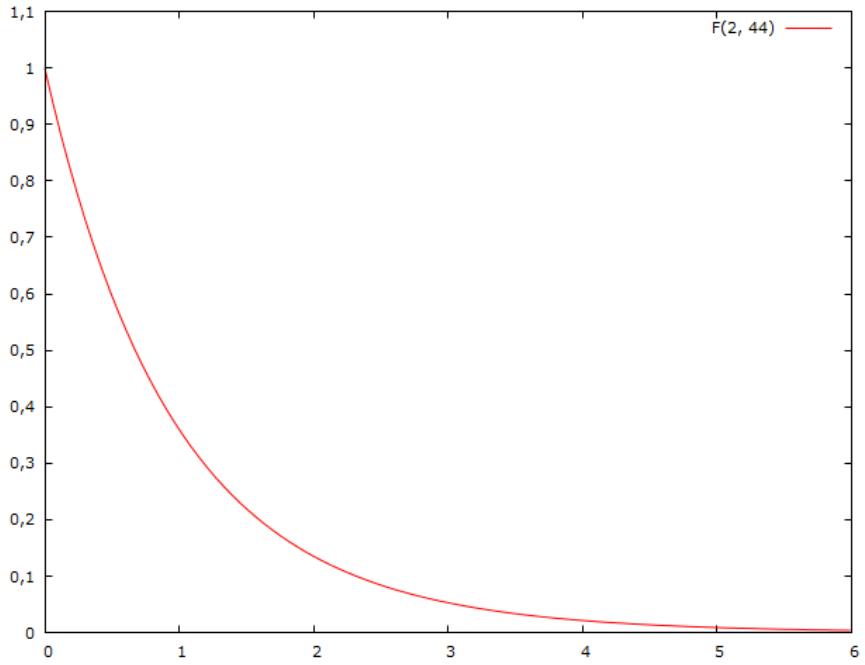



Рисунок 2.21 – распределение вероятности тестовой статистики в случае, когда спецификация модели адекватна

### 2.5.3 Тест Фишера

Нулевая гипотеза – модель в целом не значима

Таблица 2.2 – тест Фишера

| | |
|---|---|
| t-стат | 49,36391 |
| p-знач | $1{,}15 \cdot 10^{-20}$ |
| $t_{кр}$ 1% | 2,82025 |
| решение | Значима на 1%-ом уровне значимости |

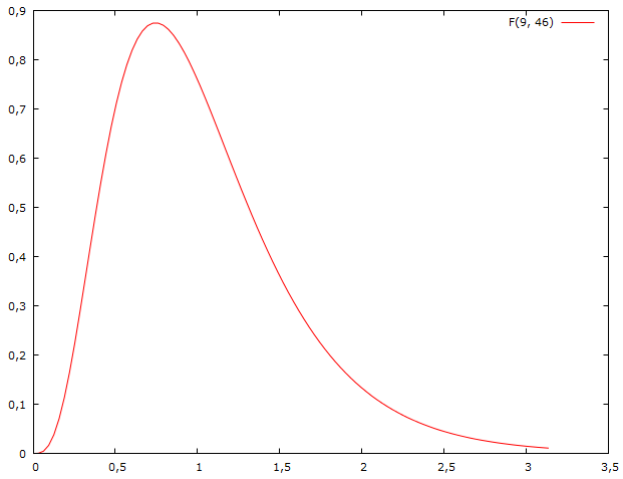



Рисунок 2.22 – распределение вероятности тестовой статистики в случае, когда модель незначима (9,46)

Модель значима на 1%-ом уровне значимости, нулевая гипотеза отвергается.

### 2.5.4 Тест Стьюдента

Нулевая гипотеза – параметр незначим.

$t_{кр}$ 10% = 1,67866, $t_{кр}$ 5% = 2,0129, $t_{кр}$ 1% = 2,68701

Таблица 2.2 – проверка значимости коэффициентов

| Регрессор | t-стат | p-знач | $t_{кр}$ | решение |
|---|---|---|---|---|
| RU | 2,519 | 0,0153 | $t_{кр}$ 5% = 2,0129 | Значим на 5% |
| EUR | 0,4510 | 0,6541 | $t_{кр}$ 10% = 1,67866 | Не значим |
| PURAN | 2,849 | 0,0065 | $t_{кр}$ 1% = 2,68701 | Значим на 1% |
| EXOIL | 1,965 | 0,00650 | $t_{кр}$ 10% = 1,67866 | Значим на 10% |
| PCRUD | -0,02669 | 0,9788 | $t_{кр}$ 10% = 1,67866 | Не значим |
| PGAS | -5,080 | $6{,}71 \cdot 10^{-6}$ | $t_{кр}$ 1% = 2,68701 | Значим на 1% |
| POIL | 2,277 | 0,0275 | $t_{кр}$ 5% = 2,0129 | Значим на 5% |
| EENER | 1,297 | 0,2010 | $t_{кр}$ 10% = 1,67866 | Не значим |
| EMP | -6,748 | $2{,}18 \cdot 10^{-8}$ | $t_{кр}$ 1% = 2,68701 | Значим на 1% |

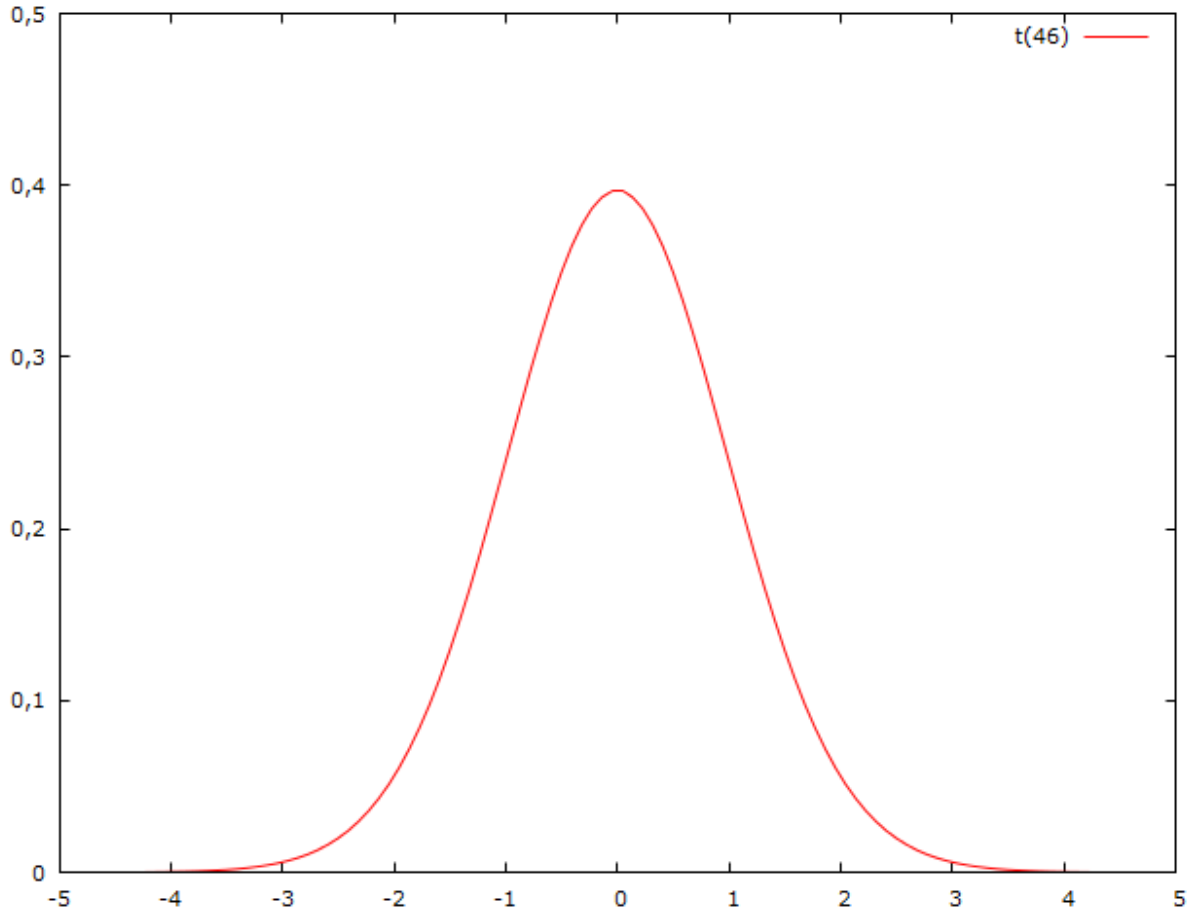



Рисунок 2.23 – распределение вероятности тестовой статистики в случае, когда коэффициент незначим (48 степеней свободы)

PURAN, EMP, PGAS значимы 1%-ом уровне значимости.

RU, POIL значимы на 5%-ом уровне значимости.

EXOIL значим на 1%-ом уровне значимости.

EUR, EENER и PCRUD не значимы.

### 2.5.5 Анализ мультиколлинеарности параметров с помощью факторов инфляции дисперсии

Строится новая модель, исключающая исследуемый параметр XXX.

```
Метод инфляционных факторов
Минимальное возможное значение = 1.0
Значения > 10.0 могут указывать на наличие мультиколлинеарности

          RU    1,888
         EUR    1,760
       PURAN   10,730
       EXOIL    7,518
       PCRUD    5,294
        PGAS    4,032
        POIL    5,759
       EENER    1,754
         EMP    9,855

VIF(j) = 1/(1 - R(j)^2), где R(j) - это коэффициент множественной корреляции
между переменной j и другими независимыми переменными

Belsley-Kuh-Welsch collinearity diagnostics:
```

Рисунок 2.24 – метод факторов инфляции дисперсии

Фактор инфляции дисперсии для PURAN больше 10. Наибольшее значение за исключением PURAN имеет EMP.

```
Анализ главных компонент
n = 56

Собственные значения для матрицы корреляций

Компонента  Собс. знач.  Доля   Интегральная
    1       1,9010       0,9505       0,9505
    2       0,0990       0,0495       1,0000

Собственные векторы (нагрузка на компоненты)

                PC1      PC2
PURAN         0,707   -0,707
EMP           0,707    0,707
```

Рисунок 2.25 – анализ главных параметров

Строим новую модель, где регрессоры PURAN и EMP заменены на собственные векторы PC1 и PC2.

```
Модель 2: МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: XXX

              Коэффициент      Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const      10584,4          1653,59            6,401      7,24e-08   ***
  RU            31,1444         12,3630          2,519      0,0153     **
  EUR          612,657        1358,34            0,4510     0,6541
  EXOIL          0,000135275     6,88526e-05     1,965      0,0555     *
  PCRUD         -2,58548e-06     9,68836e-05    -0,02669    0,9788
  PGAS        -270,711          53,2875         -5,080      6,71e-06   ***
  POIL          26,3846         11,5867          2,277      0,0275     **
  EENER          0,00501844      0,00386844      1,297      0,2010
  PC1         -387,976          48,4042         -8,015      2,81e-010  ***
  PC2         -999,038         204,396          -4,888      1,28e-05   ***

Среднее зав. перемен    14365,57   Ст. откл. зав. перемен  1025,097
Сумма кв. остатков       5422632   Ст. ошибка модели       343,3414
R-квадрат               0,906175   Испр. R-квадрат         0,887818
F(9, 46)                49,36391   P-значение (F)          1,15e-20
Лог. правдоподобие     -400,9213   Крит. Акаике            821,8426
Крит. Шварца            842,0961   Крит. Хеннана-Куинна    829,6948
Параметр rho            0,472851   Стат. Дарбина-Вотсона   1,038005

Исключая константу, наибольшее p-значение получено для переменной 6 (PCRUD)

Тест Рамсея (RESET) -
  Нулевая гипотеза: спецификация адекватна
  Тестовая статистика: F(2, 44) = 3,2791
  p-значение = P(F(2, 44) > 3,2791) = 0,0470481
```

Рисунок 2.26 – модель МНК с собственными векторами PC1 и PC2 (PURAN и EMP заменены)

```
Метод инфляционных факторов
Минимальное возможное значение = 1.0
Значения > 10.0 могут указывать на наличие мультиколлинеарности

          RU    1,888
         EUR    1,760
       EXOIL    7,518
       PCRUD    5,294
        PGAS    4,032
        POIL    5,759
       EENER    1,754
         PC1    2,078
         PC2    1,929

VIF(j) = 1/(1 - R(j)^2), где R(j) - это коэффициент множественной корреляции
между переменной j и другими независимыми переменными

Belsley-Kuh-Welsch collinearity diagnostics:

  variance proportions
```

Рисунок 2.27 –факторы инфляции дисперсии для новой модели

### 2.5.6 Проверка на избыточные переменные

```
Модель 2: МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: XXX

              Коэффициент      Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const      10584,4          1653,59            6,401      7,24e-08   ***
  RU            31,1444         12,3630          2,519      0,0153     **
  EUR          612,657        1358,34            0,4510     0,6541
  EXOIL          0,000135275     6,88526e-05     1,965      0,0555     *
  PCRUD         -2,58548e-06     9,68836e-05    -0,02669    0,9788
  PGAS        -270,711          53,2875         -5,080      6,71e-06   ***
  POIL          26,3846         11,5867          2,277      0,0275     **
  EENER          0,00501844      0,00386844      1,297      0,2010
  PC1         -387,976          48,4042         -8,015      2,81e-010  ***
  PC2         -999,038         204,396          -4,888      1,28e-05   ***

Среднее зав. перемен     14365,57   Ст. откл. зав. перемен  1025,097
Сумма кв. остатков        5422632   Ст. ошибка модели       343,3414
R-квадрат                0,906175   Испр. R-квадрат         0,887818
F(9, 46)                 49,36391   P-значение (F)          1,15e-20
Лог. правдоподобие      -400,9213   Крит. Акаике            821,8426
Крит. Шварца             842,0961   Крит. Хеннана-Куинна    829,6948
Параметр rho             0,472851   Стат. Дарбина-Вотсона   1,038005

Исключая константу, наибольшее p-значение получено для переменной 6 (PCRUD)

Тест Рамсея (RESET) -
  Нулевая гипотеза: спецификация адекватна
  Тестовая статистика: F(2, 44) = 3,2791
  p-значение = P(F(2, 44) > 3,2791) = 0,0470481
```

Рисунок 2.28а – «длинная» модель МНК

```
Модель 3: МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: XXX

              Коэффициент      Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const      11116,7          887,660           12,52      6,91e-017  ***
  RU            35,8477         11,0872          3,233      0,0022     ***
  EXOIL          0,000167899     5,32734e-05     3,152      0,0028     ***
  PGAS        -251,187          40,0764         -6,268      9,04e-08   ***
  POIL          22,0702         10,8787          2,029      0,0479     **
  PC1         -391,014          43,6559         -8,957      6,84e-012  ***
  PC2         -973,379         196,743          -4,947      9,26e-06   ***

Среднее зав. перемен     14365,57   Ст. откл. зав. перемен  1025,097
Сумма кв. остатков        5664679   Ст. ошибка модели       340,0084
R-квадрат                0,901987   Испр. R-квадрат         0,889986
F(6, 49)                 75,15572   P-значение (F)          5,36e-23
Лог. правдоподобие      -402,1440   Крит. Акаике            818,2880
Крит. Шварца             832,4655   Крит. Хеннана-Куинна    823,7846
Параметр rho             0,507482   Стат. Дарбина-Вотсона   0,980284

Тест Рамсея (RESET) -
  Нулевая гипотеза: спецификация адекватна
  Тестовая статистика: F(2, 47) = 2,68151
  p-значение = P(F(2, 47) > 2,68151) = 0,0789257
```

Рисунок 2.28б – «короткая» модель МНК

```
Тестирование модели 2:

  Нулевая гипотеза: параметры регрессии нулевые
    EUR, PCRUD, EENER
  Тестовая статистика: F(3, 46) = 0,684426, P-значение 0,566135
  Omitting variables improved 3 of 3 information criteria.

Модель 4: МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: XXX

              Коэффициент     Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const     11116,7          887,660           12,52      6,91e-017  ***
  RU           35,8477        11,0872           3,233     0,0022     ***
  EXOIL         0,000167899    5,32734e-05      3,152     0,0028     ***
  PGAS       -251,187         40,0764          -6,268     9,04e-08   ***
  POIL         22,0702        10,8787           2,029     0,0479     **
  PC1        -391,014         43,6559          -8,957     6,84e-012  ***
  PC2        -973,379        196,743           -4,947     9,26e-06   ***

Среднее зав. перемен    14365,57   Ст. откл. зав. перемен  1025,097
Сумма кв. остатков       5664679   Ст. ошибка модели       340,0084
R-квадрат               0,901987   Испр. R-квадрат         0,889986
F(6, 49)                75,15572   P-значение (F)          5,36e-23
Лог. правдоподобие     -402,1440   Крит. Акаике            818,2880
Крит. Шварца            832,4655   Крит. Хеннана-Куинна    823,7846
Параметр rho            0,507482   Стат. Дарбина-Вотсона   0,980284
```

Рисунок 2.28в – тест на избыточные переменные

$H_0$ – параметры EUR, PCRUD и EENER являются незначимыми. $t_{кр}(10\%) = 2{,}20689$. t-стат = 0,684426. P-значение = 0,566135. Значит нулевая гипотеза не отвергается, параметры EUR, PCRUD и EENER являются незначимыми на 10%-ом уровне значимости. Исключаем EUR, PCRUD и EENER

```
F(3, 46)
 Правосторонняя вероятность = 0,1
 Дополняющая вероятность = 0,9

 Критическое значение = 2,20689
```

Рисунок 2.29 – Расчет критического значения

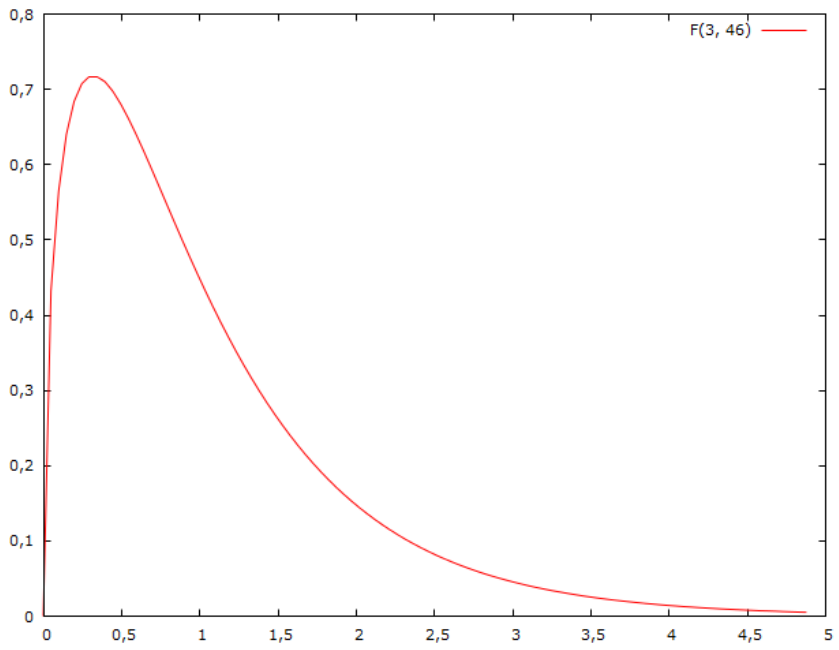



Рисунок 2.30 – Распределение F-статистики

Теперь необходимо провести RESET тест для «короткой модели»

```
Вспомогательная регрессия для теста Рамсея
МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: XXX

             Коэффициент        Ст. ошибка     t-статистика  P-значение
  ---------------------------------------------------------------------
  const      85984,9          275334              0,3123       0,7562
  RU           368,219          1509,50           0,2439       0,8083
  EXOIL          0,00167025       0,00707961      0,2359       0,8145
  PGAS       -2612,27          10576,3           -0,2470       0,8060
  POIL         224,860           926,619          0,2427       0,8093
  PC1        -4121,58          16545,1           -0,2491       0,8044
  PC2       -10114,1           41021,4           -0,2466       0,8063
  yhat^2        -0,000872529      0,00304655     -0,2864       0,7758
  yhat^3         2,55105e-08      7,31079e-08     0,3489       0,7287

Тестовая статистика: F = 2,681513,
p-значение = P(F(2,47) > 2,68151) = 0,0789
```

Рисунок 2.31 – тест Рамсея (квадраты и кубы)

Нулевая гипотеза – спецификация модели адекватна (квадраты и кубы). $t_{кр}$(5%) = 3,19506. t-стат = 2,68151. P-значение = 0,0789257. Значит нулевая гипотеза не отвергается, спецификация модели адекватна на 5%-ом уровне значимости.

## 3. Продолжение анализа модели

### 3.1 Тест Феррара-Глобера

```
Коэффициенты корреляции, наблюдения 2014:12 - 2019:07
5% критические значения (двухсторонние) = 0,2632 для n = 56

             RU          EXOIL           PGAS           POIL            PC1
         1,0000        -0,0954        -0,2565        -0,4002         0,1617  RU
                        1,0000         0,4513         0,8195        -0,3289  EXOIL
                                       1,0000         0,4912         0,2440  PGAS
                                                      1,0000        -0,3947  POIL
                                                                     1,0000  PC1

            PC2
        -0,0665  RU
        -0,5489  EXOIL
        -0,5158  PGAS
        -0,3666  POIL
         0,0000  PC1
         1,0000  PC2
```

Рисунок 3.1 – коэффициенты корреляция

| | RU | EXOIL | PGAS | POIL | PC1 | PC2 |
|---|---|---|---|---|---|---|
| RU | 1 | -0,0954 | -0,2565 | -0,4002 | 0,1617 | -0,0665 |
| EXOIL | -0,0954 | 1 | 0,4513 | 0,8195 | -0,3289 | -0,5489 |
| PGAS | -0,2565 | 0,4513 | 1 | 0,4912 | 0,244 | -0,5158 |
| POIL | -0,4002 | 0,8195 | 0,4912 | 1 | -0,3947 | -0,3666 |
| PC1 | 0,1617 | -0,3289 | 0,244 | -0,3947 | 1 | 0 |
| PC2 | -0,0665 | -0,5489 | -0,5158 | -0,3666 | 0 | 1 |

Рисунок 3.2 – Корреляционная матрица

| | |
|---|---|
| Определитель, p | 0,052267 |
| T | 56 |
| m | 6 |
| FG | 153,9641 |

Рисунок 3.3 – Расчет тестовой статистики

```
Хи-квадрат(15)
 Правосторонняя вероятность = 0,01
 Дополняющая вероятность = 0,99

 Критическое значение = 30,5779
```

Рисунок 3.3 – Поиск $t_{кр}$

```
Хи-квадрат(15): площадь правее 153,964 = 3,92897e-025
(левее: 1)
```

Рисунок 3.4 – Поиск p-значения

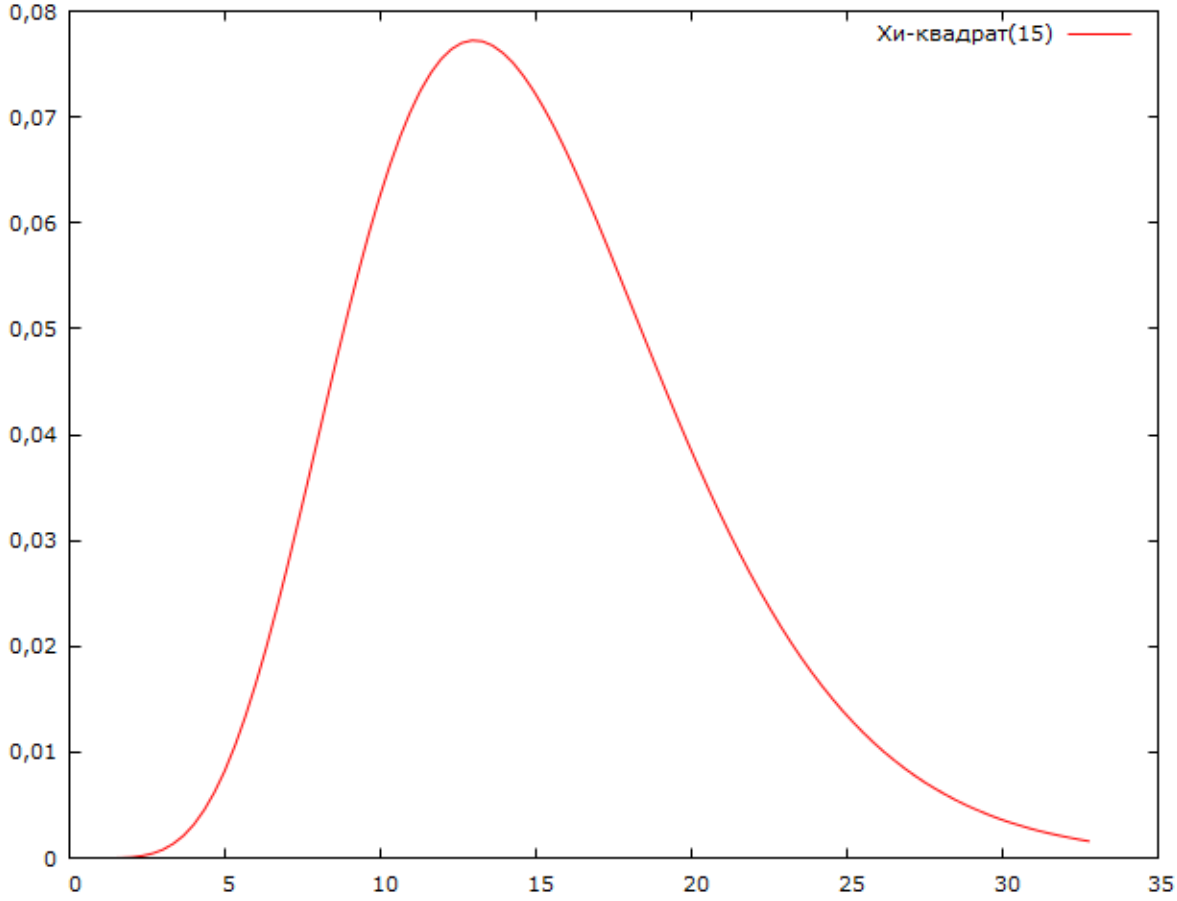



Рисунок 3.5 – Распределение тестовой статистики

Нулевая гипотеза – мультиколлинеарность отсутствует. $t_{кр}(1\%) = 30{,}5779$. t-стат = 153,9641. P-значение = $3{,}92897 \cdot 0^{-25}$. Значит нулевая гипотеза не отвергается, мультиколлинеарность присутствует на 1%-ом уровне значимости, тест не пройден.

## 3.2 Тесты по проверке гомоскедастичности

### 3.2.1 Тест Уайта

```
Модель 3: МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: XXX

              Коэффициент     Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const      11116,7          887,660           12,52      6,91e-017  ***
  RU            35,8477        11,0872           3,233     0,0022     ***
  EXOIL          0,000167899    5,32734e-05      3,152     0,0028     ***
  PGAS        -251,187         40,0764          -6,268     9,04e-08   ***
  POIL          22,0702        10,8787           2,029     0,0479     **
  PC1         -391,014         43,6559          -8,957     6,84e-012  ***
  PC2         -973,379        196,743           -4,947     9,26e-06   ***

Среднее зав. перемен    14365,57   Ст. откл. зав. перемен  1025,097
Сумма кв. остатков       5664679   Ст. ошибка модели       340,0084
R-квадрат               0,901987   Испр. R-квадрат         0,889986
F(6, 49)                75,15572   P-значение (F)          5,36e-23
Лог. правдоподобие     -402,1440   Крит. Акаике            818,2880
Крит. Шварца            832,4655   Крит. Хеннана-Куинна    823,7846
Параметр rho            0,507482   Стат. Дарбина-Вотсона   0,980284

Тест Рамсея (RESET) -
  Нулевая гипотеза: спецификация адекватна
  Тестовая статистика: F(2, 47) = 2,68151
  p-значение = P(F(2, 47) > 2,68151) = 0,0789257

Тест Вайта (White) на гетероскедастичность -
  Нулевая гипотеза: гетероскедастичность отсутствует
  Тестовая статистика: LM = 41,281
  p-значение = P(Хи-квадрат(27) > 41,281) = 0,0386937

Тест Бриша-Пэгана (Breusch-Pagan) на гетероскедастичность -
  Нулевая гипотеза: гетероскедастичность отсутствует
  Тестовая статистика: LM = 13,7184
  p-значение = P(Хи-квадрат(6) > 13,7184) = 0,0329454

Тест на нормальное распределение ошибок -
  Нулевая гипотеза: ошибки распределены по нормальному закону
  Тестовая статистика: Хи-квадрат(2) = 1,7133
  p-значение = 0,424582
```

Рисунок 3.6 – модель МНК

```
Тест Вайта (White) на гетероскедастичность
МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: uhat^2

              Коэффициент        Ст. ошибка     t-статистика  P-значение
  ---------------------------------------------------------------------
  const          -8,18270e+06       8,43038e+06   -0,9706        0,3400
  RU         148893            213282              0,6981        0,4909
  EXOIL          -0,354995          0,721729      -0,4919        0,6266
  PGAS      -923897            520580             -1,775         0,0868   *
  POIL       297387            151400              1,964         0,0595   *
  PC1        211065            784044              0,2692        0,7897
  PC2        984668                 3,16184e+06    0,3114        0,7578
  sq_RU        -502,365          1310,51          -0,3833        0,7044
  X2_X3           0,0106940         0,00963540     1,110         0,2765
  X2_X4        7206,58           6857,39           1,051         0,3023
  X2_X5       -4133,80           1921,74          -2,151         0,0402   **
  X2_X6        -894,992          8205,90          -0,1091        0,9139
  X2_X7      -15050,4           40769,2           -0,3692        0,7148
  sq_EXOIL       -7,44071e-08       2,35899e-08   -3,154         0,0038   ***
  X3_X4           0,00757364        0,0211057      0,3588        0,7224
  X3_X5           0,0175762         0,00912208     1,927         0,0642   *
  X3_X6           5,42631e-05       0,0502349      0,001080      0,9991
  X3_X7          -0,0200548         0,113511      -0,1767        0,8610
  sq_PGAS     50919,0           15426,2            3,301         0,0026   ***
  X4_X5       -5913,04           4989,44          -1,185         0,2459
  X4_X6      -35075,4           29155,6           -1,203         0,2390
  X4_X7       28668,2           82637,1            0,3469        0,7312
  sq_POIL     -1386,83           1062,45          -1,305         0,2024
  X5_X6         764,410          6532,38           0,1170        0,9077
  X5_X7       -6642,44          28527,4           -0,2328        0,8176
  sq_PC1      33025,3           48087,3            0,6868        0,4979
  X6_X7      -48710,1          134937             -0,3610        0,7208
  sq_PC2     149044            256931              0,5801        0,5665

  Неисправленный R-квадрат = 0,737161

Тестовая статистика: TR^2 = 41,281027,
p-значение = P(Хи-квадрат(27) > 41,281027) = 0,038694
```

Рисунок 3.7 – тест Уайта

Нулевая гипотеза – гомоскедастичность присутствует. $t_{кр}(1\%)$ = 46,9629. t-стат = 41,281027. P-значение = 0,038694. Значит нулевая гипотеза не отвергается, гомоскедастичность присутствует на 1%-ом уровне значимости.

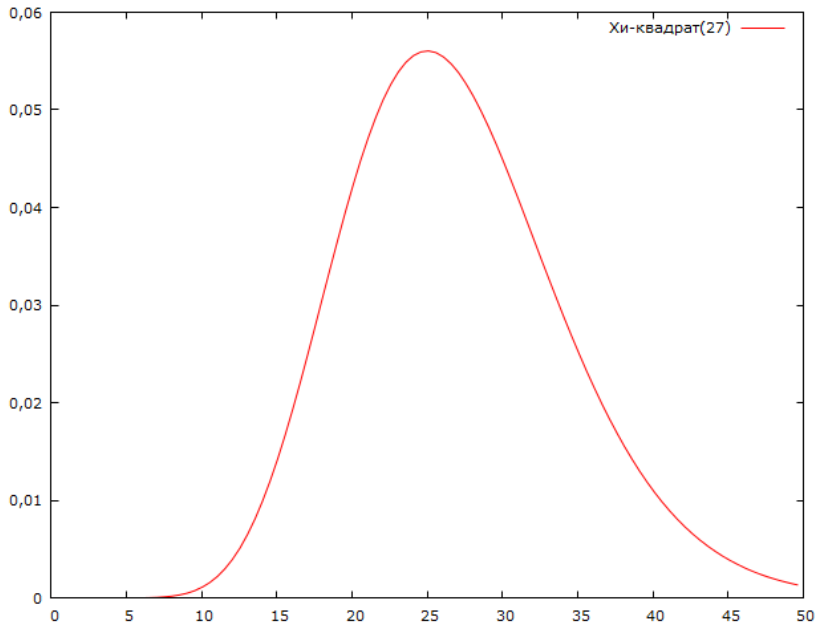



Рисунок 3.8 – тестовая статистика для теста Уайта

### 3.2.2 Тест Бриша-Пэгана

```
Тест Бриша-Пэгана (Breusch-Pagan) на гетероскедастичность
МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: Масштабированное uhat^2

             Коэффициент    Ст. ошибка    t-статистика   P-значение
  ---------------------------------------------------------------
  const       0,128544      3,35160          0,03835       0,9696
  RU          0,0124570     0,0418628        0,2976        0,7673
  EXOIL       4,39348e-07   2,01148e-07      2,184         0,0338   **
  PGAS       -0,375640      0,151319        -2,482         0,0165   **
  POIL       -0,0238674     0,0410753       -0,5811        0,5639
  PC1         0,0821170     0,164835         0,4982        0,6206
  PC2        -0,471195      0,742857        -0,6343        0,5288

  Объясненная сумма квадратов = 27,4367

Тестовая статистика: LM = 13,718362,
p-значение = P(Хи-квадрат(6) > 13,718362) = 0,032945
```

Рисунок 3.9 – тест Бриша-Пэгана

Нулевая гипотеза – гомоскедастичность присутствует. $t_{кр}$(1%) = 16,8119. t-стат = 13,748362. P-значение = 0,032945. Значит нулевая гипотеза не отвергается, гомоскедастичность присутствует на 1%-ом уровне значимости.

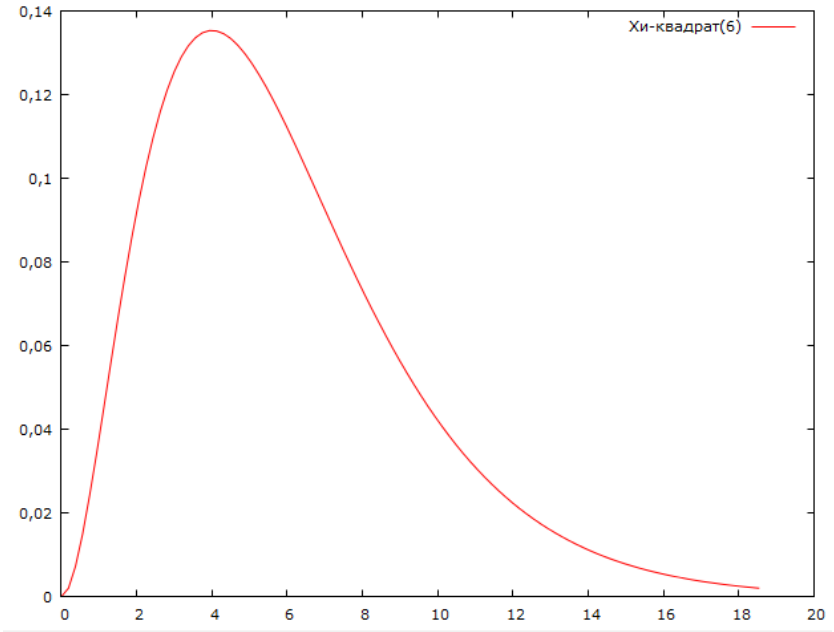



Рисунок 3.10 – тестовая статистика для теста Бриша-Пэгана

### 3.3. Тест на нормальное распределение ошибок

```
Распределение частот для uhat3, наблюдения 1-56
Количество столбцов = 7, среднее = -1,55913e-012, ст. откл. = 340,008

       интервал          середина   частота    отн.     инт.

            < -526,83   -647,71        3      5,36%    5,36% *
   -526,83 - -285,07   -405,95        6     10,71%   16,07% ***
   -285,07 - -43,317   -164,19       19     33,93%   50,00% ************
   -43,317 -  198,44    77,561       14     25,00%   75,00% *********
    198,44 -  440,19    319,32        8     14,29%   89,29% *****
    440,19 -  681,95    561,07        4      7,14%   96,43% **
          >=  681,95    802,83        2      3,57%  100,00% *

Нулевая гипотеза - нормальное распределение:
Хи-квадрат(2) = 1,713 p-значение 0,42458
```

Рисунок 3.11 – тест на нормальное распределение ошибок

Нулевая гипотеза – ошибки распределены по нормальному закону. $t_{кр}$(10%) = 4,60517. t-стат = 1,7133. P-значение = 0,424582. Значит нулевая гипотеза не отвергается, ошибки распределены по нормальному закону на 10%-ом уровне значимости.

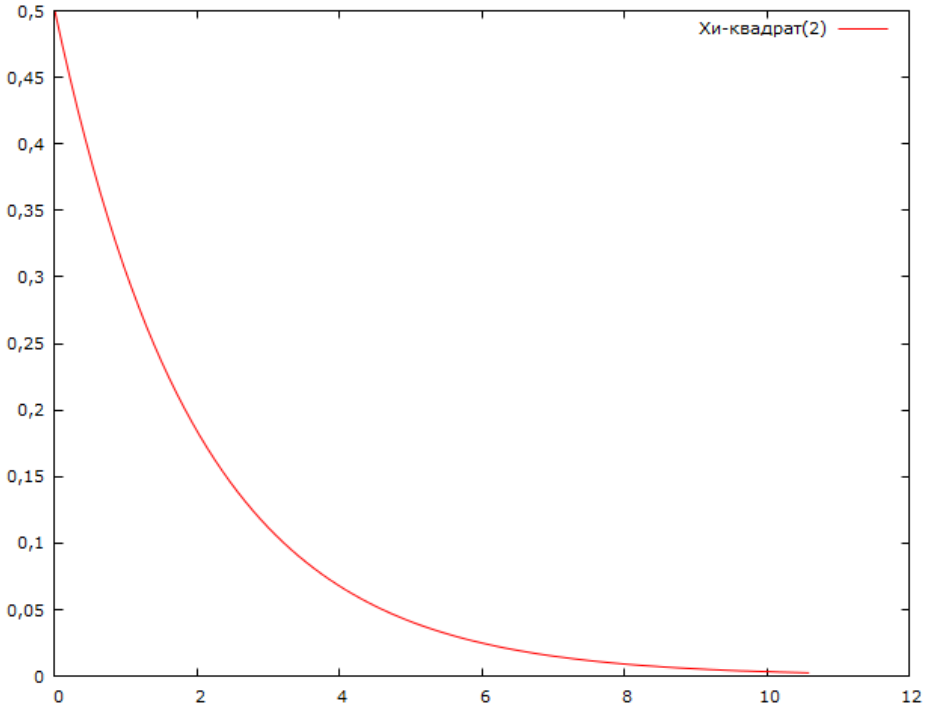



Рисунок 3.12 – распределение тестовой статистики теста на нормальность распределения ошибок

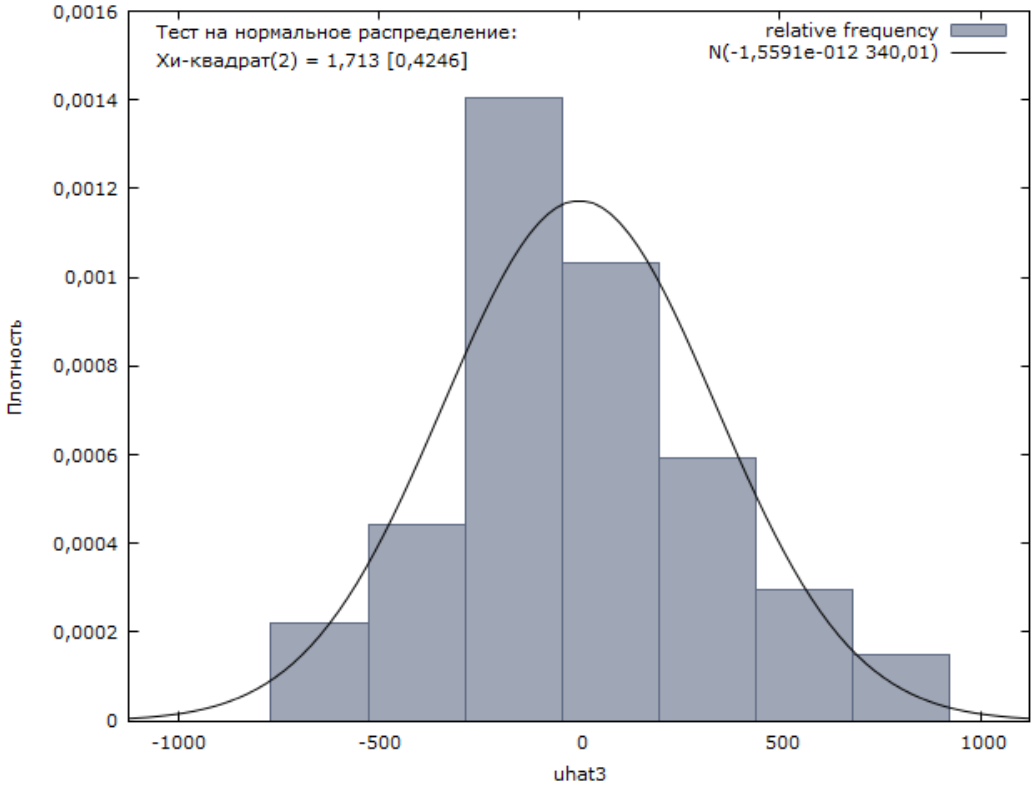



Рисунок 3.13 – плотность распределения ошибок

```
Модель 3: МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: XXX

              Коэффициент     Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const      11116,7          887,660           12,52      6,91e-017  ***
  RU            35,8477        11,0872           3,233     0,0022     ***
  EXOIL          0,000167899    5,32734e-05      3,152     0,0028     ***
  PGAS        -251,187         40,0764          -6,268     9,04e-08   ***
  POIL          22,0702        10,8787           2,029     0,0479     **
  PC1         -391,014         43,6559          -8,957     6,84e-012  ***
  PC2         -973,379        196,743           -4,947     9,26e-06   ***

Среднее зав. перемен    14365,57   Ст. откл. зав. перемен  1025,097
Сумма кв. остатков       5664679   Ст. ошибка модели       340,0084
R-квадрат               0,901987   Испр. R-квадрат         0,889986
F(6, 49)                75,15572   P-значение (F)          5,36e-23
Лог. правдоподобие     -402,1440   Крит. Акаике            818,2880
Крит. Шварца            832,4655   Крит. Хеннана-Куинна    823,7846
Параметр rho            0,507482   Стат. Дарбина-Вотсона   0,980284

Тест Рамсея (RESET) -
  Нулевая гипотеза: спецификация адекватна
  Тестовая статистика: F(2, 47) = 2,68151
  p-значение = P(F(2, 47) > 2,68151) = 0,0789257

Тест Вайта (White) на гетероскедастичность -
  Нулевая гипотеза: гетероскедастичность отсутствует
  Тестовая статистика: LM = 41,281
  p-значение = P(Хи-квадрат(27) > 41,281) = 0,0386937

Тест Бриша-Пэгана (Breusch-Pagan) на гетероскедастичность -
  Нулевая гипотеза: гетероскедастичность отсутствует
  Тестовая статистика: LM = 13,7184
  p-значение = P(Хи-квадрат(6) > 13,7184) = 0,0329454

Тест на нормальное распределение ошибок -
  Нулевая гипотеза: ошибки распределены по нормальному закону
  Тестовая статистика: Хи-квадрат(2) = 1,7133
  p-значение = 0,424582
```

Рисунок 3.14 – модель МНК

### 3.3. Вывод уравнения

| | Среднее | Медиана | S.D. | Min | Max |
|---|---|---|---|---|---|
| PURAN | 27,51 | 25,87 | 6,134 | 18,57 | 39,32 |
| EMP | 162,7 | 154,1 | 20,76 | 139,7 | 200,6 |

Рисунок 3.15 – Описательная статистика

| | PC1 | PC2 | | Среднее | S.D. | | | |
|---|---|---|---|---|---|---|---|---|
| PURAN | 0,707 | -0,707 | | 27,51 | 6,134 | | | |
| EMP | 0,707 | 0,707 | | 162,7 | 20,76 | | | |
| | | | | | | | | |
| **11116,7** | **-391,01** | **-973,38** | | | | | | |
| PURAN | -45,068 | 112,191 | | PURAN | 67,1229 | | | |
| EMP | -13,3163 | -33,1493 | | EMP | -46,4656 | | | |
| | | | | | | | | |
| | | | | Const | **16830,1** | | | |
| PURAN | 1239,82 | -3086,37 | | | | | | |
| EMP | 2166,57 | 5393,39 | | | | | | |
| | | | | | | | | |
| Итогвое уравнение имеет вид: | | | | | | | | |
| Y= 16830-35,85*RU-0,000168*EXOIL-251,19*PGAS+22,07*POIL+67,12*PURAN-46,47*EMP | | | | | | | | |

Рисунок 3.15 – Вывод уравнения

Выводы:

1. При увеличении курса национальной валюты по отношению к доллару США на 1 рубль стоимость фиксированного набора потребительских товаров и услуг падает на 35,85 рублей.

2. При увеличении мировой цены на уран на 1 доллар за фунт стоимость фиксированного набора потребительских товаров и услуг увеличится на 67,12 рублей

3. При увеличении экспорта сырой нефти из Российской Федерации на 1 млрд долларов стоимость фиксированного набора потребительских товаров и услуг уменьшится на 168 рублей.

4. При увеличении цены на природный газ на 1 доллар за кубометр стоимость фиксированного набора потребительских товаров и услуг падает на 251,19 рублей.

5. При увеличении цен на сырую нефть: Западный Техас, Средний уровень (WTI) на 1$ за баррель стоимость фиксированного набора потребительских товаров и услуг возрастет на 22,07 рублей.

6. При увеличении всех занятых человек в добыче нефти и газа в США на 1 тысячу человек стоимость фиксированного набора потребительских товаров и услуг падает на 46,47 рублей.

### 3.3. Построение фиктивного прогноза

```
Модель 2: МНК, использованы наблюдения 2014:12-2019:06 (T = 55)
Зависимая переменная: XXX

              Коэффициент     Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const     11032,1          860,713           12,82      4,13e-017  ***
  RU           37,7500        10,7781           3,502     0,0010     ***
  EXOIL         0,000140331    5,33077e-05      2,632     0,0114     **
  PGAS       -225,175         40,8214          -5,516     1,36e-06   ***
  POIL         22,4871        10,5383           2,134     0,0380     **
  PC1        -408,462         43,1239          -9,472     1,45e-012  ***
  PC2       -1001,31         191,036           -5,241     3,53e-06   ***

Среднее зав. перемен    14333,64   Ст. откл. зав. перемен  1006,041
Сумма кв. остатков       5205373   Ст. ошибка модели       329,3103
R-квадрат               0,904758   Испр. R-квадрат         0,892853
F(6, 48)                75,99687   P-значение (F)          8,33e-23
Лог. правдоподобие     -393,1330   Крит. Акаике            800,2660
Крит. Шварца            814,3173   Крит. Хеннана-Куинна    805,6998
Параметр rho            0,570984   Стат. Дарбина-Вотсона   0,872951
```

Рисунок 3.16 – Модель с отсутствием крайнего временного лага

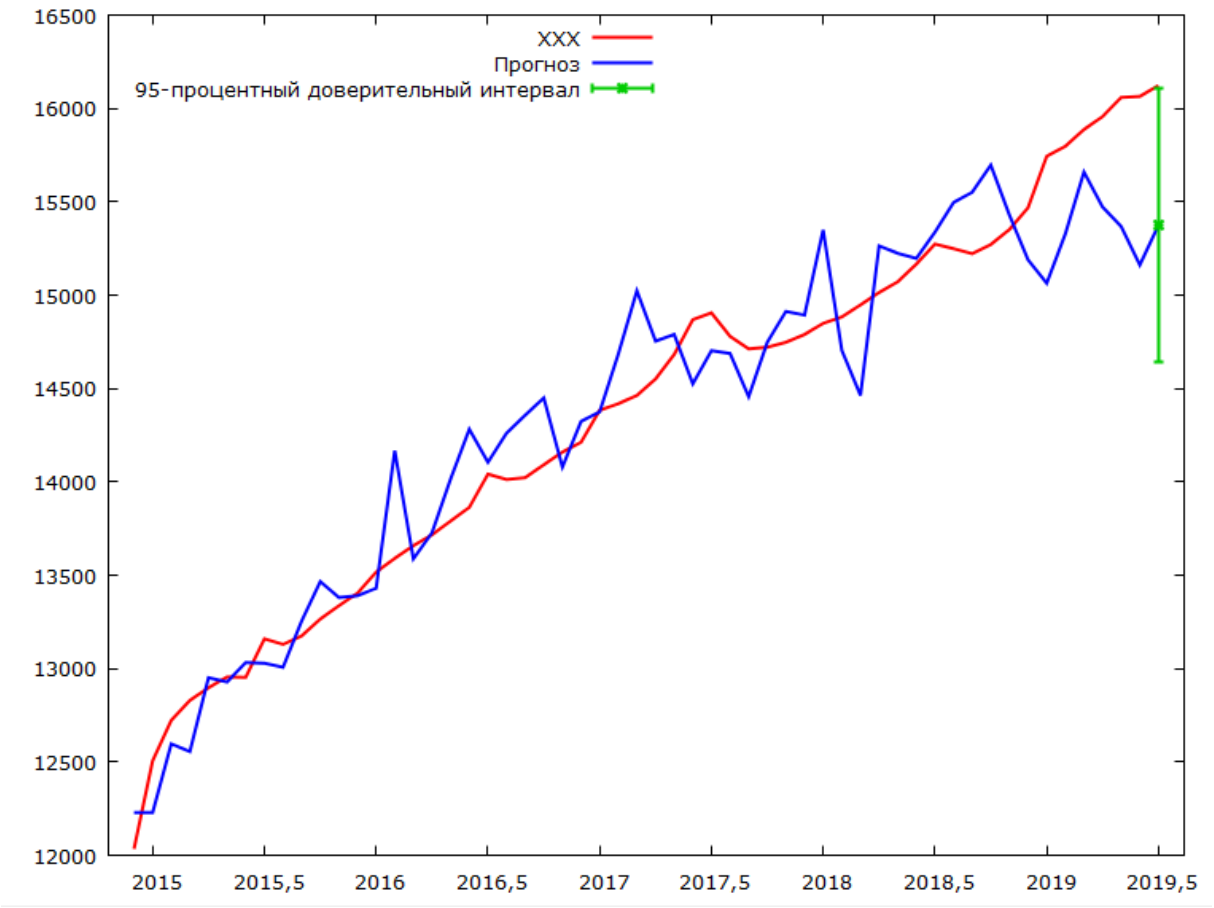



Рисунок 3.17 – Прогноз

```
Статистика для оценки прогноза using 1 observations

Средняя ошибка (ME)                                747,64
Корень из средней квадратичной ошибки (RMSE)       747,64
Средняя абсолютная ошибка (MAE)                    747,64
Средняя процентная ошибка (MPE)                    4,6374
Средняя абсолютная процентная ошибка (MAPE)        4,6374
U-статистика Тейла (Theil's U)                     0
```

Рисунок 3.18 – Статистика для оценки прогноза

Последнее значение попало в 95%-ый доверительный интервал.

Вывод: Прогноз подтвердился

### 3.4 Автокорреляция остатков

#### 3.4.1 Тест Дарбина-Уотсона

Нулевая гипотеза – автокорреляция отсутствует.

Модель 1

```
Модель 1: МНК, использованы наблюдения 2014:12-2019:07 (T = 56)
Зависимая переменная: XXX

              Коэффициент     Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const     11116,7          887,660           12,52      6,91e-017  ***
  RU           35,8477        11,0872           3,233     0,0022     ***
  EXOIL         0,000167899    5,32734e-05      3,152     0,0028     ***
  PGAS       -251,187         40,0764          -6,268     9,04e-08   ***
  POIL         22,0702        10,8787           2,029     0,0479     **
  PC1        -391,014         43,6559          -8,957     6,84e-012  ***
  PC2        -973,379        196,743           -4,947     9,26e-06   ***

Среднее зав. перемен    14365,57   Ст. откл. зав. перемен  1025,097
Сумма кв. остатков       5664679   Ст. ошибка модели       340,0084
R-квадрат               0,901987   Испр. R-квадрат         0,889986
F(6, 49)                75,15572   P-значение (F)          5,36e-23
Лог. правдоподобие     -402,1440   Крит. Акаике            818,2880
Крит. Шварца            832,4655   Крит. Хеннана-Куинна    823,7846
Параметр rho            0,507482   Стат. Дарбина-Вотсона   0,980284
```

Рисунок 3.19 – Модель МНК

```
5% критические значения для статистики Дарбина-Вотсона, n = 56, k = 6

  dL = 1,3424
  dU = 1,8124
```

Рисунок 3.20 – Критические значения

DW = 0.980284

0<DW<DL, значит Но отвергается, положительная автокорреляция присутствует.

DW = 2(1- ρ)

ρ = 0.509858

### 3.4.2 Избавление от автокорреляции

```
Модель 1: МНК, использованы наблюдения 2015:01-2019:07 (T = 55)
Зависимая переменная: XXX_NEW

              Коэффициент     Ст. ошибка    t-статистика  P-значение
  ---------------------------------------------------------------------
  const       5540,65         496,189          11,17      6,05e-015  ***
  RU_NEW        37,9850        12,3941          3,065     0,0036     ***
  EXOIL_NEW      5,78808e-05    3,45139e-05     1,677     0,1000
  PGAS_NEW    -144,857         47,1861         -3,070     0,0035     ***
  POIL_NEW      21,6021         9,03382         2,391     0,0208     **
  PC2_NEW     -748,007        210,696          -3,550     0,0009     ***
  PC_NEW      -465,334         59,8864         -7,770     4,91e-010  ***

Среднее зав. перемен    7099,828   Ст. откл. зав. перемен  474,7166
Сумма кв. остатков       3028525   Ст. ошибка модели       251,1857
R-квадрат               0,751132   Испр. R-квадрат         0,720024
F(6, 48)                24,14560   P-значение (F)          6,00e-13
Лог. правдоподобие     -378,2386   Крит. Акаике            770,4772
Крит. Шварца            784,5285   Крит. Хеннана-Куинна    775,9109
Параметр rho            0,581251   Стат. Дарбина-Вотсона   0,838473

Исключая константу, наибольшее p-значение получено для переменной 22 (EXOIL_NEW)
```

Рисунок 3.21 – Модель МНК

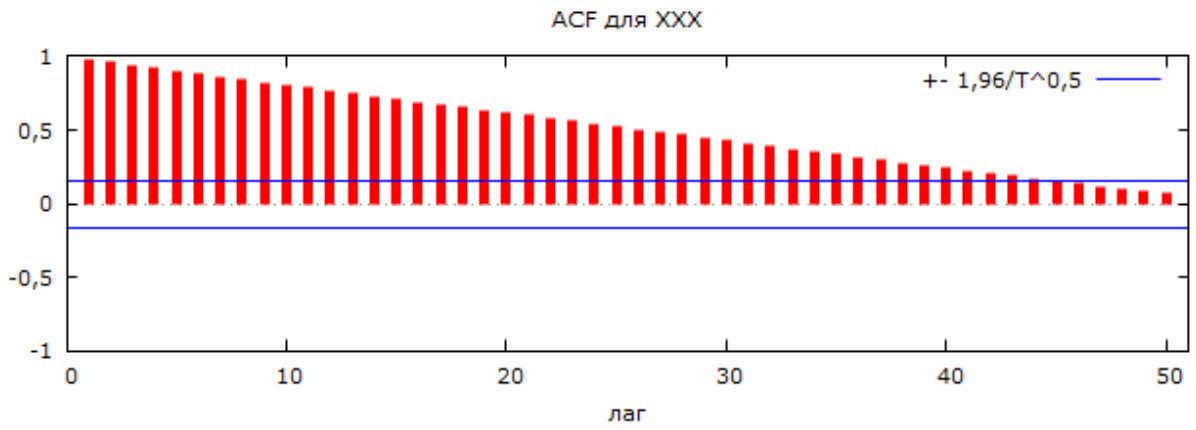
```
5% критические значения для статистики Дарбина-Вотсона, n = 55, k = 6

  dL = 1,3344
  dU = 1,8137
```

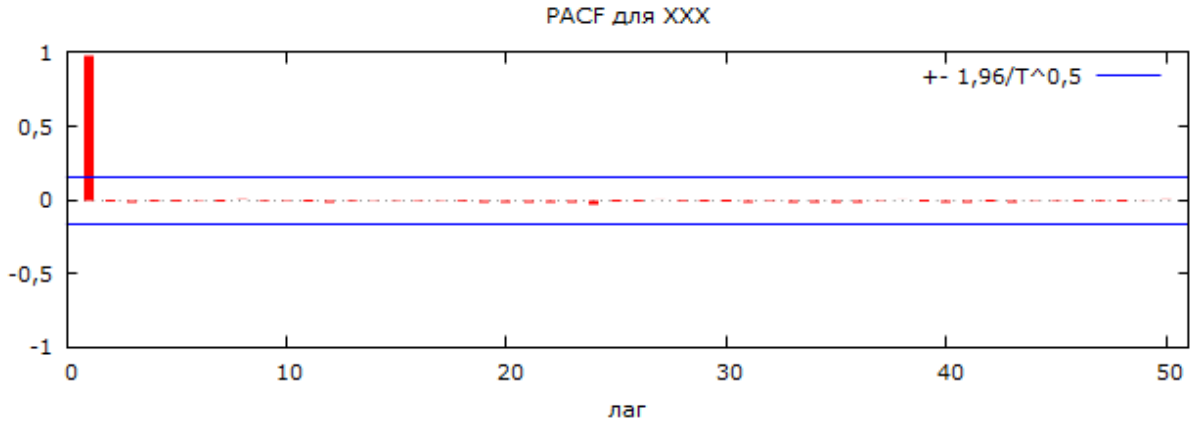
Рисунок 3.22 – Критические значения

DW = 0,838473

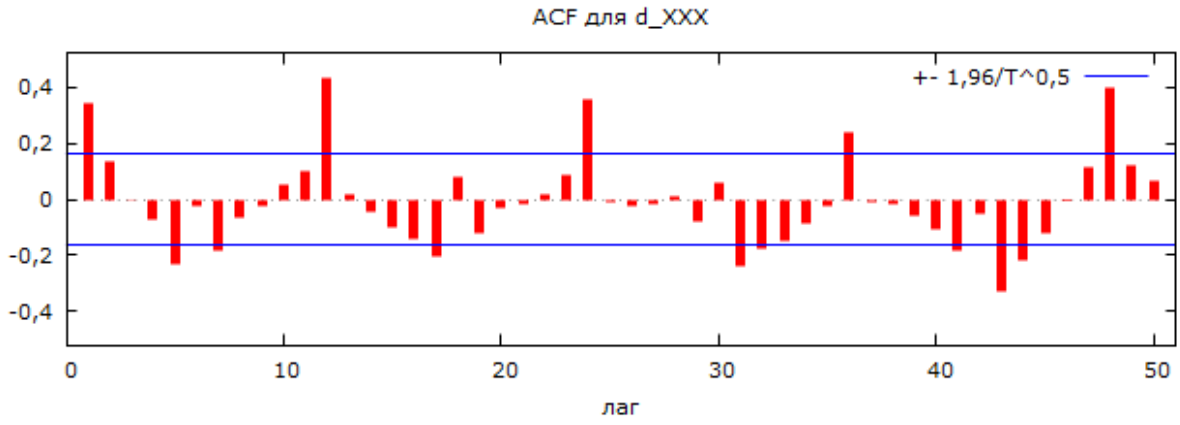
0<DW<DL, значит Но отвергается, положительная автокорреляция присутствует.

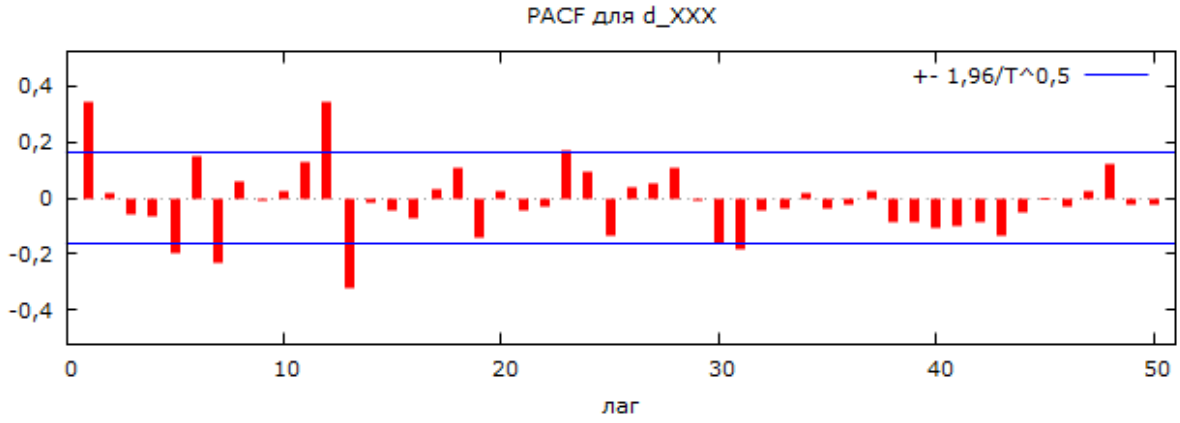
## 4.1 Коррелограмма




Так как на коррелограмме ряда первая автокорреляция приближается к единице, то приводим коррелограмму для временного ряда первой разности.




У данного временного ряда присутствует сезонная составляющая и есть подозрение на наличие единичного корня (стохастического тренда)

### 4.2 тест Дики-Фуллера

```
Расширенный тест Дики-Фуллера для XXX
testing down from 13 lags, criterion Крит. Акаике
объем выборки 137
нулевая гипотеза единичного корня: a = 1

  тест без константы
  включая 13 лага(-ов) для (1-L)XXX
  модель: (1-L)y = (a-1)*y(-1) + ... + e
  оценка для (a - 1): 0,00242218
  тестовая статистика: tau_nc(1) = 2,10343
  асимпт. p-значение 0,992
  коэф. автокорреляции 1-го порядка для e: 0,013
  лаг для разностей: F(13, 123) = 8,219 [0,0000]
```

тест "без константы" Нулевая гипотеза – есть единичный корень. tстат = 2,10343 P-значение = 0,992

Значит нулевая гипотеза не отвергается на всех стандартных уровнях значимости, единичный корень присутствует.

```
Расширенный тест Дики-Фуллера для XXX
testing down from 13 lags, criterion Крит. Акаике
объем выборки 137
нулевая гипотеза единичного корня: a = 1

  тест с константой
  включая 13 лага(-ов) для (1-L)XXX
  модель: (1-L)y = b0 + (a-1)*y(-1) + ... + e
  оценка для (a - 1): -0,000234156
  тестовая статистика: tau_c(1) = -0,126552
  асимпт. p-значение 0,9448
  коэф. автокорреляции 1-го порядка для e: 0,004
  лаг для разностей: F(13, 122) = 7,876 [0,0000]
```

тест "с константой" Нулевая гипотеза – есть единичный корень tкр(10%) = -2,57 t-стат = -0,126552 P-значение = 0,9448

Значит нулевая гипотеза не отвергается на всех стандартных уровнях значимости, единичный корень присутствует.

```
Расширенный тест Дики-Фуллера для XXX
testing down from 13 lags, criterion Крит. Акаике
объем выборки 137
нулевая гипотеза единичного корня: a = 1

  с константой и трендом
  включая 13 лага(-ов) для (1-L)XXX
  модель: (1-L)y = b0 + b1*t + (a-1)*y(-1) + ... + e
  оценка для (a - 1): -0,0409552
  тестовая статистика: tau_ct(1) = -2,05262
  асимпт. p-значение 0,5717
  коэф. автокорреляции 1-го порядка для e: 0,016
  лаг для разностей: F(13, 121) = 8,081 [0,0000]
```

тест "с константой и трендом" Нулевая гипотеза – есть единичный корень tкр(10%) = -3,12 t-стат = -2,05262 P-значение = 0,5717

Значит нулевая гипотеза не отвергается на всех стандартных уровнях значимости, единичный корень присутствует

```
Расширенный тест Дики-Фуллера для XXX
testing down from 13 lags, criterion Крит. Акаике
объем выборки 137
нулевая гипотеза единичного корня: a = 1

  with constant, linear and quadratic trend
  включая 13 лага(-ов) для (1-L)XXX
  модель: (1-L)y = b0 + b1*t + b2*t^2 + (a-1)*y(-1) + ... + e
  оценка для (a - 1): -0,0470635
  тестовая статистика: tau_ctt(1) = -1,9975
  асимпт. p-значение 0,8181
  коэф. автокорреляции 1-го порядка для e: 0,018
  лаг для разностей: F(13, 120) = 8,039 [0,0000]
```

тест "с константой, трендом и квадратичным трендом" Нулевая гипотеза – есть единичный корень t-стат = -1,9975 P-значение = 0,818

Значит нулевая гипотеза не отвергается на всех стандартных уровнях значимости, единичный корень присутствует. Также стоит отметить, что минимальное значение тестовой статистики (с учетом знака) наблюдается на тесте "с константой и трендом". Значит делаем вывод о присутствии детерминированного тренда.

### 4.3 Выбор параметров и построение модели ARIMA

| AR (сезонная) | I | AR | MA | AIC | BIC |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 1527,485 | 1539,165 |
| 1 | 1 | 2 | 0 | 1518,861 | 1533,425 |
| 1 | 1 | 3 | 0 | 1527,208 | 1516,86 |
| 1 | 1 | 4 | 0 | 1501,495 | 1521,78 |
| 1 | 1 | 5 | 0 | 1491,422 | 1514,545 |
| 1 | 1 | 6 | 0 | 1482,95 | 1508,895 |
| 1 | 1 | 7 | 0 | 1474,694 | 1503,446 |
| 1 | 1 | 8 | 0 | 1466,482 | 1498,025 |
| 1 | 1 | 9 | 0 | 1458,297 | 1492,614 |
| 1 | 1 | 10 | 0 | **1450,115** | **1487,192** |
| | | | | | |
| | | | | | |
| AR (сезонная) | I | AR | MA | AIC | BIC |
| 1 | 1 | 10 | 0 | 1450,115 | 1487,192 |
| 1 | 1 | 10 | 1 | 1448,78 | 1491,56 |
| 1 | 1 | 10 | 2 | **1439,297** | **1484,929** |
| 1 | 1 | 10 | 3 | 1449,831 | 1498,316 |
| 1 | 1 | 10 | 4 | 1452,074 | 1503,411 |
| 1 | 1 | 10 | 5 | 1458,154 | 1480,172 |
| 1 | 1 | 10 | 6 | 1453,94 | 1477,116 |
| 1 | 1 | 10 | 7 | 1454,854 | 1479,188 |
| 1 | 1 | 10 | 8 | 1456,564 | 1482,057 |
| 1 | 1 | 10 | 9 | 1459,669 | 1486,322 |
| 1 | 1 | 10 | 10 | 1460,869 | 1488,68 |

По критерию Акаике и Шварца наилучший порядок авторегрессии AR = 10, скользящего среднего по критерию Акаике и Шварца MA = 2.

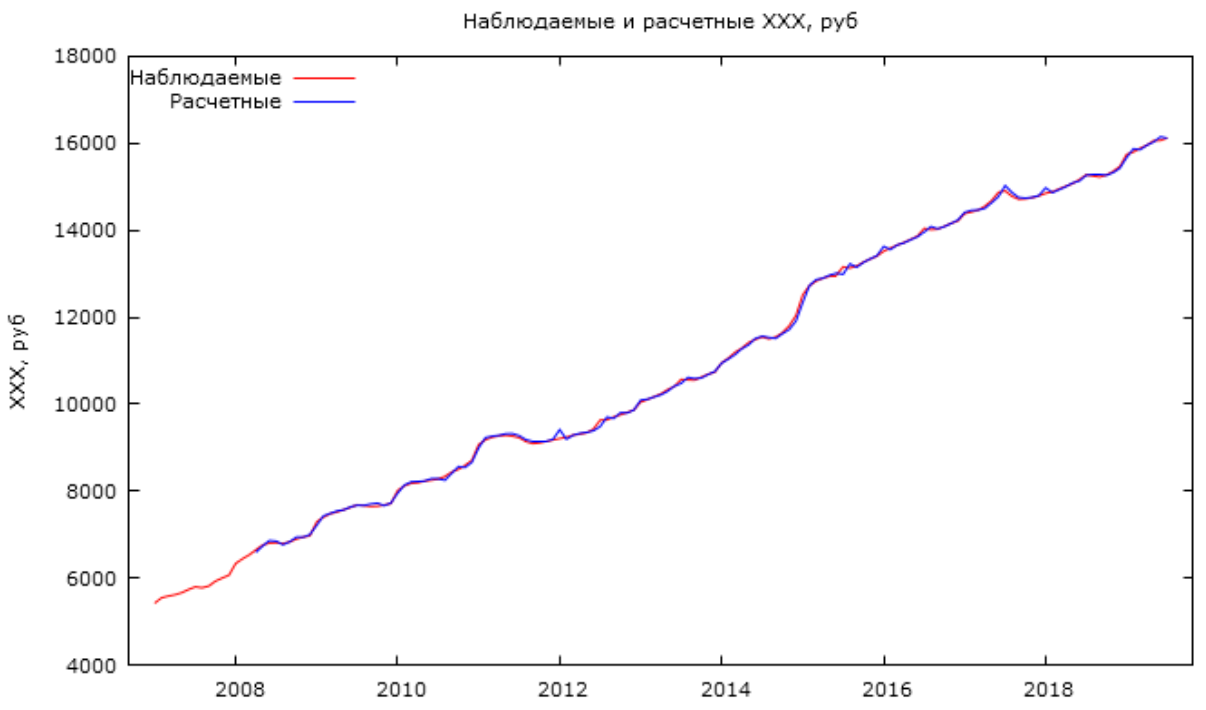
```
Оценок функции: 113
Оценок градиента: 34

Модель 7: ARMAX, использованы наблюдения 2008:12-2019:07 (T = 128)
Оценено с помощью метода BHHH (Berndt, Hall, Hall and Hausman) (условный метод МП)
Зависимая переменная: (1-L) XXX

             Коэффициент   Ст. ошибка      z       P-значение
  ---------------------------------------------------------
  const       5,72211      5,02604        1,138    0,2549
  phi_1       2,07995      0,0914838     22,74     1,99e-114  ***
  phi_2      -1,67871      0,265037      -6,334    2,39e-010  ***
  phi_3       0,420672     0,443789       0,9479   0,3432
  phi_4       0,114062     0,474979       0,2401   0,8102
  phi_5      -0,274644     0,397687      -0,6906   0,4898
  phi_6       0,251982     0,310025       0,8128   0,4163
  phi_7      -0,0551077    0,247552      -0,2226   0,8238
  phi_8      -0,0551402    0,302969      -0,1820   0,8556
  phi_9       0,0486745    0,306773       0,1587   0,8739
  phi_10     -0,0665115    0,145319      -0,4577   0,6472
  Phi_1       0,686674     0,0634246     10,83     2,57e-027  ***
  theta_1    -1,66999      0,0595345    -28,05     3,91e-173  ***
  theta_2     0,999998     0,0466700     21,43     7,48e-102  ***
  time       -0,0105817    0,0451000     -0,2346   0,8145

Среднее зав. перемен    71,70836   Ст. откл. зав. перемен  82,72020
Среднее инноваций      -0,930417   Ст. откл. инноваций     59,04589
R-квадрат               0,999541   Испр. R-квадрат         0,999489
Лог. правдоподобие     -703,6485   Крит. Акаике            1439,297
Крит. Шварца            1484,929   Крит. Хеннана-Куинна    1457,838

               Действ. часть  Мним. часть  Модуль   Частота
  ---------------------------------------------------------
  AR
Корень  1           1,0431    -0,3523     1,1009    -0,0518
Корень  2           1,0431     0,3523     1,1009     0,0518
Корень  3           0,3994    -1,4263     1,4811    -0,2065
Корень  4           0,3994     1,4263     1,4811     0,2065
Корень  5          -1,4911     0,4611     1,5608     0,4523
Корень  6          -1,4911    -0,4611     1,5608    -0,4523
Корень  7          -0,3946    -1,4475     1,5003    -0,2924
Корень  8          -0,3946     1,4475     1,5003     0,2924
Корень  9           0,8091    -0,6135     1,0154    -0,1033
Корень 10           0,8091     0,6135     1,0154     0,1033
  AR (сезонные)
Корень  1           1,4563     0,0000     1,4563     0,0000
  MA
Корень  1           0,8350    -0,5503     1,0000    -0,0927
Корень  2           0,8350     0,5503     1,0000     0,0927
```

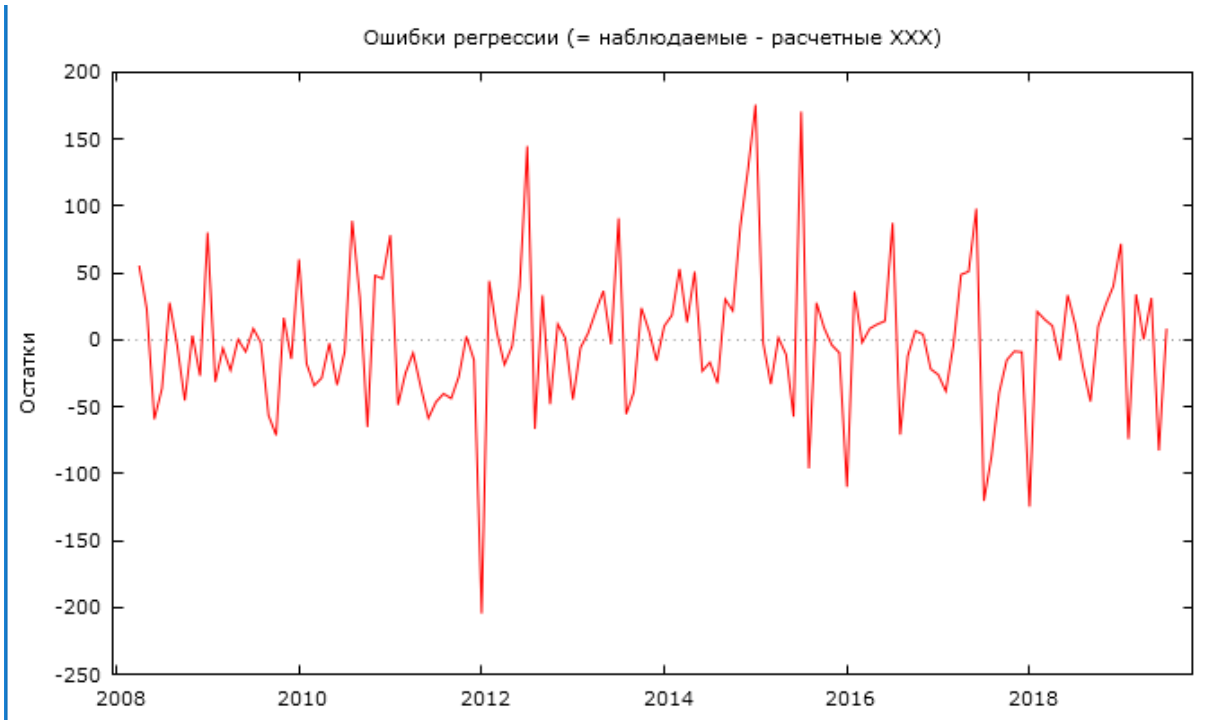
Итоговая модель – ARIMA(10,1,2) с сезонной авторегрессией первого порядка и временным трендом.

## 5.1 График наблюдаемых и расчетных значений построенной модели ARIMA; график остатков.

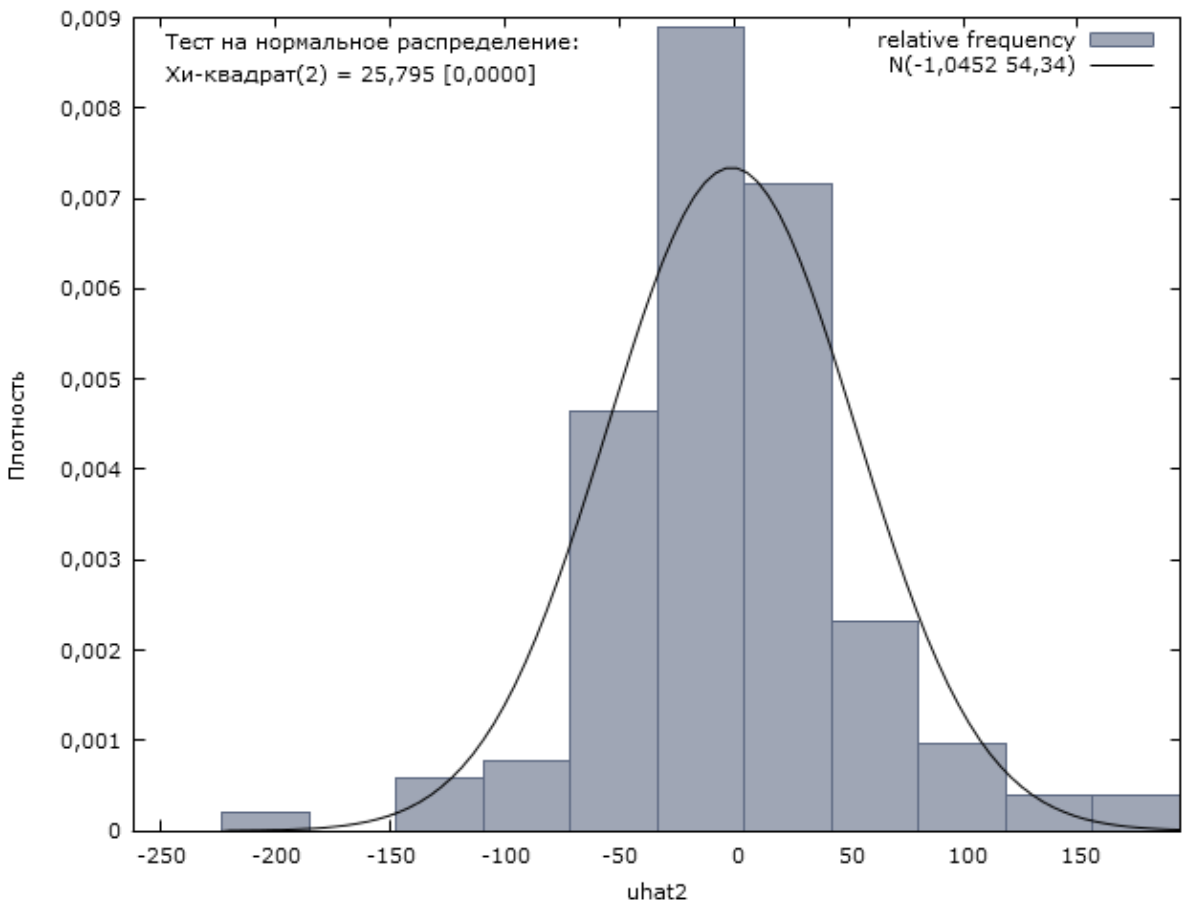
Основываясь на том, что коэффициенты AR > 2 незначимы и для AR = 2 имеется локальный минимум AIC; выберем модель с AR = 2. Проанализировав такие модели получаем, что наилучший результат будет у следующей модели: С трендом; с единичным корнем; AR = 2; MA = 0; Сезонное AR = 1; Сезонное MA = 1.

```
Оценок функции: 112
Оценок градиента: 56

Модель 1: ARMAX, использованы наблюдения 2008:04-2019:07 (T = 136)
Оценено с помощью метода BHHH (Berndt, Hall, Hall and Hausman) (условный метод МП)
Зависимая переменная: (1-L) XXX

             Коэффициент   Ст. ошибка       z       P-значение
  ---------------------------------------------------------------
  const       0,593578     4,37899        0,1356    0,8922
  phi_1       0,520636     0,0729887      7,133     9,81e-013  ***
  phi_2      -0,0233477    0,109814      -0,2126    0,8316
  Phi_1       0,961528     0,0117330     81,95      0,0000     ***
  Theta_1    -0,858948     0,0426549    -20,14      3,49e-090  ***
  time        0,00253357   0,0559017      0,04532   0,9639

Среднее зав. перемен    70,50051   Ст. откл. зав. перемен  81,05481
Среднее инноваций      -1,045153   Ст. откл. инноваций     53,13791
R-квадрат               0,999663   Испр. R-квадрат         0,999653
Лог. правдоподобие     -733,2888   Крит. Акаике            1480,578
Крит. Шварца            1500,966   Крит. Хеннана-Куинна    1488,863

               Действ. часть  Мним. часть  Модуль   Частота
  ---------------------------------------------------------------
  AR
Корень  1           2,1228     0,0000     2,1228     0,0000
Корень  2          20,1764     0,0000    20,1764     0,0000
  AR (сезонные)
Корень  1           1,0400     0,0000     1,0400     0,0000
  MA (сезонные)
Корень  1           1,1642     0,0000     1,1642     0,0000
  ---------------------------------------------------------------
```




## 5.2 Проверка нормальности распределения остатков

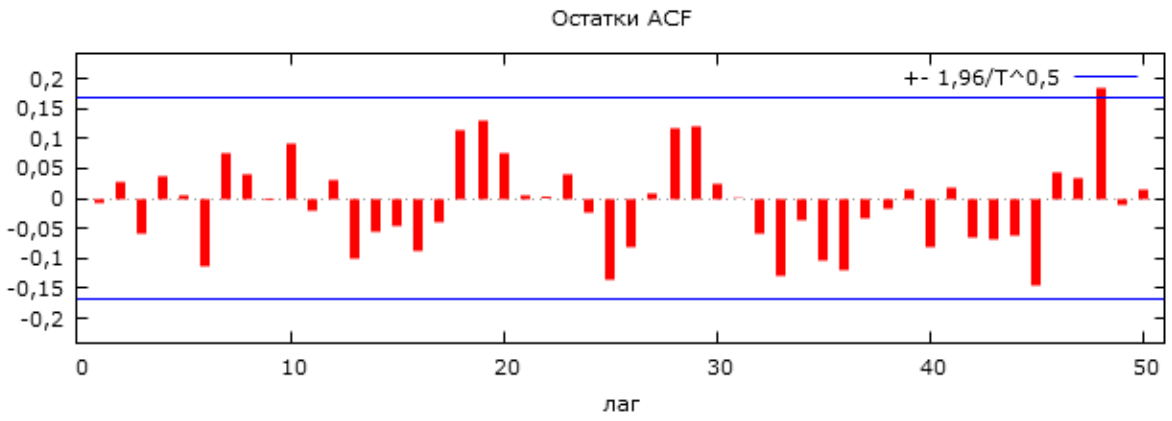


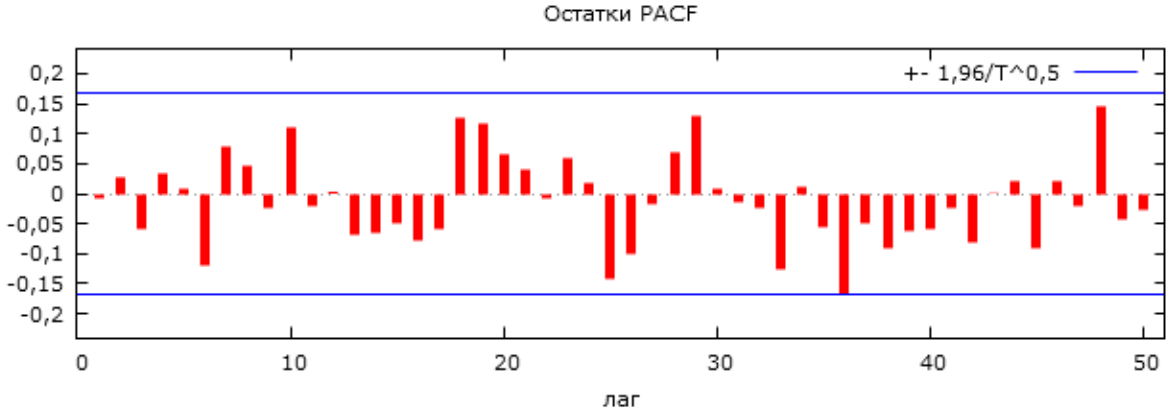

```
Распределение частот для uhat2, наблюдения 16-151
Количество столбцов = 11, среднее = -1,04515, ст. откл. = 54,3398

       интервал          середина   частота    отн.     инт.

           < -185,55   -204,57        1      0,74%    0,74%
   -185,55 - -147,52   -166,54        0      0,00%    0,74%
   -147,52 - -109,50   -128,51        3      2,21%    2,94%
   -109,50 - -71,469   -90,483        4      2,94%    5,88% *
   -71,469 - -33,441   -52,455       24     17,65%   23,53% ******
   -33,441 -  4,5868   -14,427       46     33,82%   57,35% ************
    4,5868 -  42,615    23,601       37     27,21%   84,56% *********
    42,615 -  80,642    61,628       12      8,82%   93,38% ***
    80,642 -  118,67    99,656        5      3,68%   97,06% *
    118,67 -  156,70    137,68        2      1,47%   98,53%
          >=  156,70    175,71        2      1,47%  100,00%

Нулевая гипотеза - нормальное распределение:
Хи-квадрат(2) = 25,795 p-значение 0,00000
```

На всех стандартных уровнях значимости остатки не распределены нормально

### 5.3 Проверка того, что остатки являются белым шумом

```
Оценок функции: 112
Оценок градиента: 56

Модель 1: ARMAX, использованы наблюдения 2008:04-2019:07 (T = 136)
Оценено с помощью метода BHHH (Berndt, Hall, Hall and Hausman) (условный метод МП)
Зависимая переменная: (1-L) XXX

             Коэффициент   Ст. ошибка       z       P-значение
  ---------------------------------------------------------
  const       0,593578     4,37899        0,1356    0,8922
  phi_1       0,520636     0,0729887      7,133     9,81e-013  ***
  phi_2      -0,0233477    0,109814      -0,2126    0,8316
  Phi_1       0,961528     0,0117330     81,95      0,0000     ***
  Theta_1    -0,858948     0,0426549    -20,14      3,49e-090  ***
  time        0,00253357   0,0559017      0,04532   0,9639

Среднее зав. перемен    70,50051   Ст. откл. зав. перемен  81,05481
Среднее инноваций      -1,045153   Ст. откл. инноваций     53,13791
R-квадрат               0,999663   Испр. R-квадрат         0,999653
Лог. правдоподобие     -733,2888   Крит. Акаике            1480,578
Крит. Шварца            1500,966   Крит. Хеннана-Куинна    1488,863

               Действ. часть  Мним. часть  Модуль   Частота
  ---------------------------------------------------------
  AR
Корень  1           2,1228     0,0000     2,1228     0,0000
Корень  2          20,1764     0,0000    20,1764     0,0000
  AR (сезонные)
Корень  1           1,0400     0,0000     1,0400     0,0000
  MA (сезонные)
Корень  1           1,1642     0,0000     1,1642     0,0000
  ---------------------------------------------------------
```


Остатки ACF
+- 1,96/T^0,5
0,2
0,15
0,1
0,05
0
-0,05
-0,1
-0,15
-0,2
0
10
20
30
40
50
лаг


Остатки PACF
+- 1,96/T^0,5
0,2
0,15
0,1
0,05
0
-0,05
-0,1
-0,15
-0,2
0
10
20
30
40
50
лаг

Функция автокорреляции ошибок
***, **, * indicate significance at the 1%, 5%, 10% levels
using standard error 1/T^0,5

| Лаг | ACF | PACF | | Q-стат. | [p-значение] |
|---|---|---|---|---|---|
| 1 | -0,0071 | -0,0071 | | | |
| 2 | 0,0270 | 0,0270 | | | |
| 3 | -0,0606 | -0,0602 | | | |
| 4 | 0,0357 | 0,0343 | | | |
| 5 | 0,0042 | 0,0078 | | 0,8100 | [0,368] |
| 6 | -0,1144 | -0,1206 | | 2,7008 | [0,259] |
| 7 | 0,0736 | 0,0784 | | 3,4898 | [0,322] |
| 8 | 0,0404 | 0,0474 | | 3,7287 | [0,444] |
| 9 | -0,0020 | -0,0231 | | 3,7293 | [0,589] |
| 10 | 0,0908 | 0,1103 | | 4,9578 | [0,549] |
| 11 | -0,0197 | -0,0194 | | 5,0163 | [0,658] |
| 12 | 0,0295 | 0,0031 | | 5,1481 | [0,742] |
| 13 | -0,1001 | -0,0680 | | 6,6775 | [0,671] |
| 14 | -0,0562 | -0,0670 | | 7,1633 | [0,710] |
| 15 | -0,0474 | -0,0500 | | 7,5111 | [0,756] |
| 16 | -0,0877 | -0,0780 | | 8,7137 | [0,727] |
| 17 | -0,0397 | -0,0602 | | 8,9627 | [0,776] |
| 18 | 0,1143 | 0,1250 | | 11,0410 | [0,683] |
| 19 | 0,1302 | 0,1167 | | 13,7615 | [0,544] |
| 20 | 0,0743 | 0,0665 | | 14,6557 | [0,550] |
| 21 | 0,0049 | 0,0388 | | 14,6596 | [0,620] |
| 22 | 0,0027 | -0,0083 | | 14,6608 | [0,685] |
| 23 | 0,0414 | 0,0582 | | 14,9455 | [0,726] |
| 24 | -0,0224 | 0,0164 | | 15,0298 | [0,775] |
| 25 | -0,1365 | -0,1428 | * | 18,1810 | [0,638] |
| 26 | -0,0825 | -0,1007 | | 19,3427 | [0,624] |
| 27 | 0,0090 | -0,0178 | | 19,3565 | [0,680] |
| 28 | 0,1167 | 0,0693 | | 21,7237 | [0,596] |
| 29 | 0,1215 | 0,1288 | | 24,3139 | [0,501] |
| 30 | 0,0233 | 0,0069 | | 24,4102 | [0,553] |
| 31 | 0,0009 | -0,0141 | | 24,4104 | [0,607] |
| 32 | -0,0585 | -0,0243 | | 25,0276 | [0,626] |
| 33 | -0,1300 | -0,1248 | | 28,1085 | [0,512] |

На всех стандартных уровнях значимости автокорреляция остатков отсутствует; остатки являются "белым шумом".

### 5.4 Запись уравнения построенной модели

## Точный метод

При построении модели «***точным методом***» с константой (установлен флажок «Включить константу»), коэффициенты, приведенные в распечатке модели, соответствуют зависимости:

$$X_t - \mu = a_1(X_{t-1} - \mu) + a_2(X_{t-2} - \mu) + \cdots + a_p(X_{t-p} - \mu) + \varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + \cdots + b_q\varepsilon_{t-q}$$

или (после несложных преобразований):

$$X_t = a_0 + a_1X_{t-1} + a_2X_{t-2} + \cdots + a_pX_{t-p} + \varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + \cdots + b_q\varepsilon_{t-q}$$

где

$$a_0 = \mu \cdot (1 - a_1 - a_2 - \cdots - a_p)$$

## Условный метод

При построении модели "***Условным***" методом коэффициенты, приведенные в распечатке модели, соответствуют зависимости:

$$X_t = a_0 + a_1X_{t-1} + a_2X_{t-2} + \cdots + a_pX_{t-p} + \varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + \cdots + b_q\varepsilon_{t-q}$$

## Уравнение с сезонной составляющей (месячные данные)

Пусть $a_i$ – коэффициенты AR, $A$ – коэффициент сезонной AR, $b_j$ – коэффициенты MA, $B$ – коэффициент сезонной MA.

Для ARMA без единичного корня уравнение модели записывается следующим образом:

$$(1 - a_1\hat{L} - a_2\hat{L}^2 - \cdots - a_p\hat{L}^p)(1 - A\hat{L}^{12})X_t = a_0 + \beta t + (1 + b_1\hat{L} + b_2\hat{L}^2 + \cdots + b_q\hat{L}^q)(1 + B\hat{L}^{12})\varepsilon_t$$

После раскрытия всех скобок получаем итоговое уравнение модели:

$$X_t = a_0 + \beta t + a_1X_{t-1} + a_2X_{t-2} + \cdots + a_pX_{t-p} + AX_{t-12} - Aa_1X_{t-13} - Aa_2X_{t-14} - \cdots - Aa_pX_{t-12-p} +$$

$$+\varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + \cdots + b_q\varepsilon_{t-q} + B\varepsilon_{t-12} + Bb_1\varepsilon_{t-13} + Bb_2\varepsilon_{t-14} + \cdots + Bb_q\varepsilon_{t-12-q}$$

Если модель ARIMA с единичным корнем, то вместо $X_t$ во всех слагаемых будет первая разность $\Delta X_t$.

Если модель построена «точным методом…», то константу нужно пересчитать, но учесть наличие сезонной компоненты:

$$a_0 = \mu \cdot (1 - a_1 - a_2 - \cdots - a_p) \cdot (1 - A)$$

Модель с трендом; с единичным корнем; AR = 2; MA = 0; Сезонное AR = 1; Сезонное MA = 1.

```
Оценок функции: 112
Оценок градиента: 56

Модель 1: ARMAX, использованы наблюдения 2008:04-2019:07 (T = 136)
Оценено с помощью метода BHHH (Berndt, Hall, Hall and Hausman) (условный метод МП)
Зависимая переменная: (1-L) XXX

             Коэффициент   Ст. ошибка       z       P-значение
  ---------------------------------------------------------------
  const       0,593578     4,37899        0,1356    0,8922
  phi_1       0,520636     0,0729887      7,133     9,81e-013  ***
  phi_2      -0,0233477    0,109814      -0,2126    0,8316
  Phi_1       0,961528     0,0117330     81,95      0,0000     ***
  Theta_1    -0,858948     0,0426549    -20,14      3,49e-090  ***
  time        0,00253357   0,0559017      0,04532   0,9639

Среднее зав. перемен    70,50051   Ст. откл. зав. перемен  81,05481
Среднее инноваций      -1,045153   Ст. откл. инноваций     53,13791
R-квадрат               0,999663   Испр. R-квадрат         0,999653
Лог. правдоподобие     -733,2888   Крит. Акаике            1480,578
Крит. Шварца            1500,966   Крит. Хеннана-Куинна    1488,863

               Действ. часть  Мним. часть  Модуль   Частота
  ---------------------------------------------------------------
  AR
Корень  1           2,1228      0,0000     2,1228     0,0000
Корень  2          20,1764      0,0000    20,1764     0,0000
  AR (сезонные)
Корень  1           1,0400      0,0000     1,0400     0,0000
  MA (сезонные)
Корень  1           1,1642      0,0000     1,1642     0,0000
  ---------------------------------------------------------------
```

## Уравнение с сезонной составляющей (месячные данные)

Пусть $a_i$ – коэффициенты AR, $A$ – коэффициент сезонной AR, $b_j$ – коэффициенты MA, $B$ – коэффициент сезонной MA.

Для ARMA без единичного корня уравнение модели записывается следующим образом:

$$\left(1 - a_1\hat{L} - a_2\hat{L}^2 - \cdots - a_p\hat{L}^p\right)\left(1 - A\hat{L}^{12}\right)X_t = a_0 + \beta t + \left(1 + b_1\hat{L} + b_2\hat{L}^2 + \cdots + b_q\hat{L}^q\right)\left(1 + B\hat{L}^{12}\right)\varepsilon_t$$

После раскрытия всех скобок получаем итоговое уравнение модели:

$$X_t = a_0 + \beta t + a_1X_{t-1} + a_2X_{t-2} + \cdots + a_pX_{t-p} + AX_{t-12} - Aa_1X_{t-13} - Aa_2X_{t-14} - \cdots - Aa_pX_{t-12-p} +$$

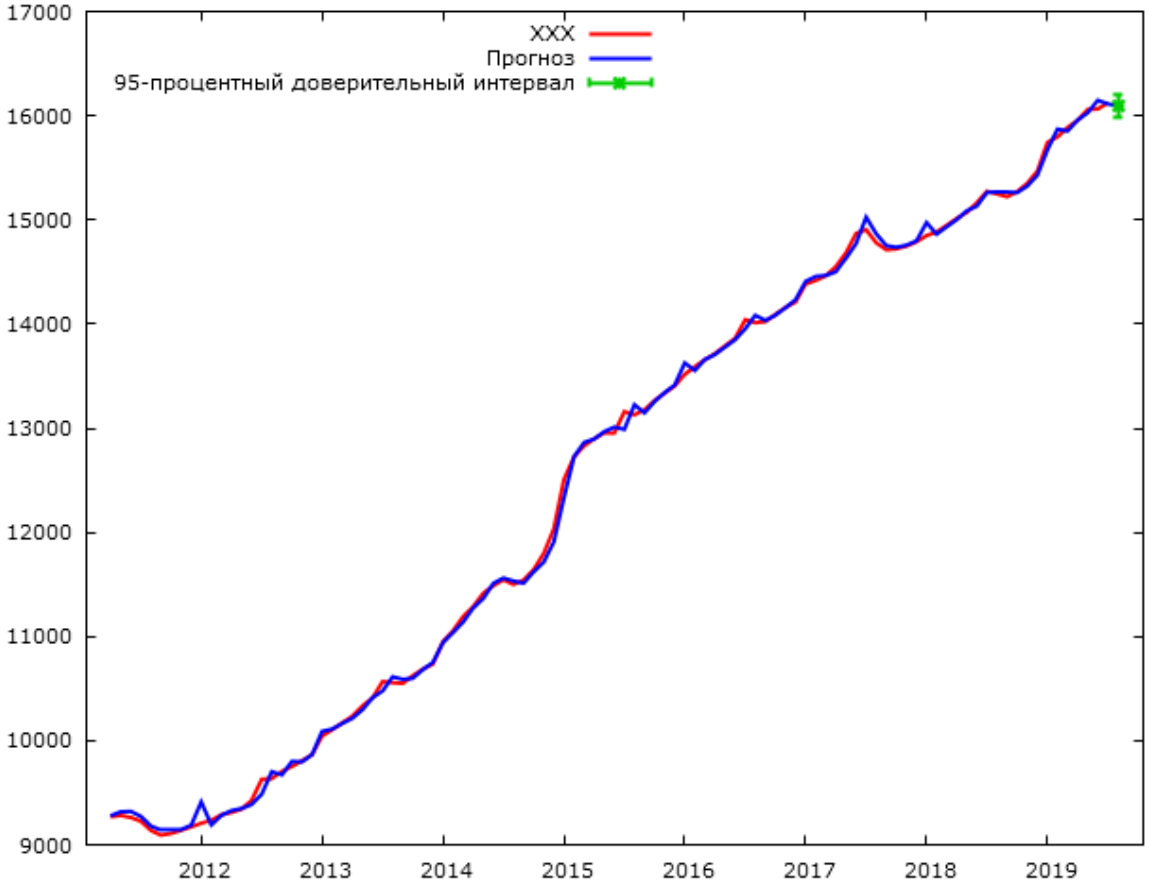
$$+\varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + \cdots + b_q\varepsilon_{t-q} + B\varepsilon_{t-12} + Bb_1\varepsilon_{t-13} + Bb_2\varepsilon_{t-14} + \cdots + Bb_q\varepsilon_{t-12-q}$$

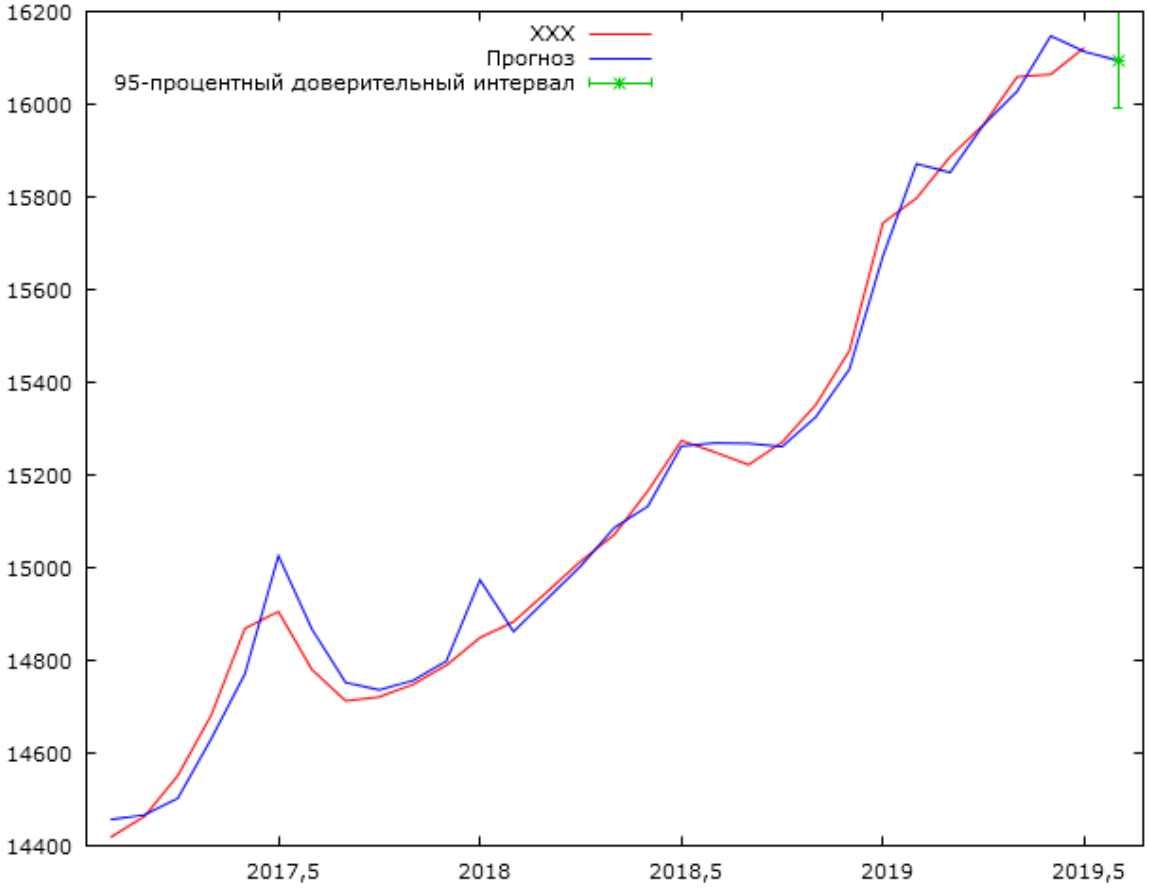
$$\Delta X_t = 0.593578 + 0.00253357t + 0.520636\Delta X_{t-1} - 0.0233477\Delta X_{t-2} + 0.961528\Delta X_{t-12} - 0.5006\Delta X_{t-13} + 0.022449\Delta X_{t-14} + \varepsilon(t) - 0.858948\varepsilon_{t-12}$$

$$Aa1 = 0.5006$$

$$Aa2 = -0.022449$$

## 5.5 Построение прогноза


XXX
Прогноз
95-процентный доверительный интервал
17000
16000
15000
14000
13000
12000
11000
10000
9000
2012
2013
2014
2015
2016
2017
2018
2019


XXX
Прогноз
95-процентный доверительный интервал
16200
16000
15800
15600
15400
15200
15000
14800
14600
14400
2017,5
2018
2018,5
2019
2019,5

```
2019:03     15886,70     15852,71
2019:04     15956,14     15955,84
2019:05     16059,24     16027,92
2019:06     16064,35     16147,22
2019:07     16121,93     16113,58
2019:08                  16094,08       53,138     15989,93 - 16198,23
```

Для временного ряда первой разности XXX построена стационарная (так как все корни AR > 1) обратимая (так как все корни MA > 1) модель с трендом ARMA(14, 12).

На основе полученной модели сделан прогноз, для которого построен 95% доверительный интервал.

### 5.6 Тест на коинтеграцию временных рядов

```
Шаг 1: тестирование единичного корня для XXX

Расширенный тест Дики-Фуллера для XXX
testing down from 4 lags, criterion Крит. Акаике
объем выборки 149
нулевая гипотеза единичного корня: a = 1

  с константой и трендом
  включая один лаг для (1-L)XXX
  модель: (1-L)y = b0 + b1*t + (a-1)*y(-1) + ... + e
  оценка для (a - 1): -0,0454216
  тестовая статистика: tau_ct(1) = -2,31085
  асимпт. p-значение 0,4274
  коэф. автокорреляции 1-го порядка для e: -0,017

Шаг 2: тестирование единичного корня для RU

Расширенный тест Дики-Фуллера для RU
testing down from 4 lags, criterion Крит. Акаике
объем выборки 148
нулевая гипотеза единичного корня: a = 1

  с константой и трендом
  включая 2 лага(-ов) для (1-L)RU
  модель: (1-L)y = b0 + b1*t + (a-1)*y(-1) + ... + e
  оценка для (a - 1): -0,042089
  тестовая статистика: tau_ct(1) = -2,16416
  асимпт. p-значение 0,5092
  коэф. автокорреляции 1-го порядка для e: -0,020
  лаг для разностей: F(2, 143) = 33,708 [0,0000]
```

```
Шаг 3: коинтеграционная регрессия

Коинтеграционная регрессия -
МНК, использованы наблюдения 2007:01-2019:07 (T = 151)
Зависимая переменная: XXX

             Коэффициент   Ст. ошибка   t-статистика   P-значение
  ---------------------------------------------------------------
  const       4666,27      49,5254         94,22       5,18e-134  ***
  RU            31,1212     2,15558        14,44       4,82e-030  ***
  time          62,5099     0,802845       77,86       5,38e-122  ***

Среднее зав. перемен    10727,61   Ст. откл. зав. перемен  3196,568
Сумма кв. остатков       6033813   Ст. ошибка модели       201,9134
R-квадрат               0,996063   Испр. R-квадрат         0,996010
Лог. правдоподобие     -1014,228   Крит. Акаике            2034,457
Крит. Шварца            2043,508   Крит. Хеннана-Куинна    2038,134
Параметр rho            0,890801   Стат. Дарбина-Вотсона   0,215855

Шаг 4: тестирование единичного корня для uhat

Расширенный тест Дики-Фуллера для uhat
testing down from 4 lags, criterion Крит. Акаике
объем выборки 147
нулевая гипотеза единичного корня: a = 1

  тест без константы
  включая 3 лага(-ов) для (1-L)uhat
  модель: (1-L)y = (a-1)*y(-1) + ... + e
  оценка для (a - 1): -0,180937
  тестовая статистика: tau_ct(2) = -4,50688
  асимпт. p-значение 0,005441
  коэф. автокорреляции 1-го порядка для e: 0,004
  лаг для разностей: F(3, 143) = 6,663 [0,0003]

There is evidence for a cointegrating relationship if:
(a) The unit-root hypothesis is not rejected for the individual variables, and
(b) the unit-root hypothesis is rejected for the residuals (uhat) from the
    cointegrating regression.
```

На уровне значимости 1% гипотеза о наличии стохастического тренда для рядов XXX и RU не отвергается, а для ряда

$$u_t = XXX_t - 31.1212RU_t - 62.5099t - 4666.27$$

отвергается; следовательно, присутствует коинтеграционная связь между рядами XXX и RU.

Линейная комбинация рядов XXX и RU

$$Z_t = XXX_t - 31.1212RU_t$$

не содержит стохастического тренда и является стационарной относительно линейного тренда.