

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра дифференциальных уравнений и математи-  
ческой экономики

**РЕГРЕССИОННЫЙ АНАЛИЗ ВЗАИМОСВЯЗИ  
ИНВЕСТИЦИЙ И ЭКОНОМИЧЕСКОГО РОСТА**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студентки 4 курса 451 группы  
направления 38.03.05 — Бизнес-информатика

механико-математического факультета  
Филиной Екатерины Алексеевны

Научный руководитель  
доцент, к. ф.-м. н., доцент

\_\_\_\_\_

В. В. Новиков

Заведующий кафедрой  
зав.кафедрой, д. ф.-м. н.,  
профессор

\_\_\_\_\_

С. И. Дудов

Саратов 2021

**Введение.** Рост уровня жизни населения России и борьба с бедностью возможны только в условиях растущей экономики. Введение антироссийских санкций в 2014 году значительно ухудшило социально-экономическое положение страны в целом и отдельных её жителей. Основная причина стагнации российской экономики – это низкий уровень инвестиций и их неэффективное использование.

Экономический рост также существенно отличается по регионам Российской Федерации. Зачастую будучи в схожих географических, климатических, экономических и историко-этнографических условиях заметны различия в уровне развития соседних субъектов федерации.

Применение эконометрических уравнений для объяснения влияния инвестиций на экономический рост регионов позволяет не только количественно оценить их влияние на экономику, но и увидеть региональные различия их эффективности. Тем самым показываются не только явные причинно-следственные связи, но и выявляется влияние скрытых факторов, таких как административные барьеры, неправомерное использование служебного положения, нецелевое использование инвестиционных средств.

В качестве показателя экономического роста региона целесообразно использовать валовой региональный продукт (ВРП), как итоговый показатель функционирования экономики субъекта Российской Федерации. А в качестве характеристики инвестиций нами рассмотрены инвестиции в основной капитал, так именно они составляют наибольшую долю от общего объема инвестиций в нефинансовые активы (99% по данным за 2020 год).

При расчете эконометрических зависимостей экономического роста от инвестиций важны не только знание макроэкономических показателей, но и понимание методологии эконометрических построений, а также наличие программного обеспечения, с помощью которого возможно осуществление всего комплекса расчетно-графических работ.

Таким образом, высокая значимость, а также недостаточная теоретическая и практическая разработанность вопросов программного

обеспечения построения моделей зависимости экономического роста регионов России от инвестиций определяют несомненную актуальность данного исследования.

**Цель исследования:** эконометрическое моделирование связи между инвестициями и экономическим ростом.

**Задачи исследования:**

- рассмотреть существующие методы и инструменты регрессионного анализа, в частности регрессионного анализа временных рядов;
- осуществить программную реализацию алгоритма эконометрического моделирования зависимости экономического роста от инвестиций;
- провести вычислительный эксперимент и дать оценку его результатов;
- сделать экономический вывод на основе полученных моделей.

**Объект исследования:** региональные экономические системы, в которых инвестиции выступает одним из основных факторов роста.

**Предмет исследования:** программное обеспечение построения моделей экономического роста региональных экономик под влиянием инвестиций.

**Степень изученности проблемы:** теоретической основой для исследования послужили научные работы разных направлений, раскрывающие отдельные аспекты исследуемой темы. К ним относятся работы по макро- и мезоэкономике, эконометрике и основам языка R. Вопросы влияния инвестиций на региональный экономический рост рассматриваются в работах Глазьева С.Ю., Гранберга А.Г., Мау В.А., Шахбанова Р.Б. и других авторов. Эконометрическому моделированию посвящены работы Айвазяна С.А., Мхитаряна В.С., Носко В.П., Сажина Ю.В. и других ученых. Программированию на языке R посвящены работы Р. Кабакова, Н. Метлоффа, Н.Ю. Золотых, А.Н. Половинкина. Однако, на данный момент отсутствуют работы, которые совмещают все три аспекта, необходимые для программного обеспечения построения моделей экономического роста региональных экономик под влиянием инвестиций.

**Теоретическую и методологическую основу** исследования составляют методы экономической теории, теории вероятностей и математической статистики; методы анализа и моделирования временных рядов, методы программирования.

**Структура работы:** работа состоит из введения, двух глав, заключения, списка используемых источников и приложений. Первая глава «Теоретические основы анализа временных рядов» посвящена изучению теоретических аспектов эконометрического моделирования регионального экономического роста. Во второй главе «Инвестиционные модели динамики факторов экономического роста Саратовской области» осуществлена программная реализация алгоритма эконометрического моделирования зависимости экономического роста регионов от инвестиций, рассчитаны конкретные регрессионные модели, дана оценка их адекватности и экономическая интерпретация.

### **Основное содержание работы**

В первой главе рассмотрены теоретические аспекты эконометрического моделирования регионального экономического роста, в частности основные элементы временного ряда, методы оценки тенденции временного ряда, использование критерия Дарбина-Уотсона, Бройша-Годфри и Вайта для выявления автокорреляции и гетероскедастичности в остатках, методология построения моделей с распределенным лагом и моделей авторегрессии, лагов Алмона.

Важным этапом спецификации эконометрической модели является выбор конкретных показателей в качестве зависимой и независимых переменных. В частности, экономический рост регионов оценивался нами на основе валового регионального продукта, данные о котором ежегодно публикуются Федеральной службой государственной статистики (Росстатом) и её территориальными органами. Для анализа были собраны данные за 2005–2019 годы. Отсутствие квартальных данных не позволяет провести оценку сезонных колебаний, а достаточно короткий временной период – выявить

циклическую составляющую. Поэтому исследование было направлено на построение эконометрических моделей с распределённым лагом. Включение в модель как текущих значений объясняющей переменной, так и значений объясняющей и результативной переменной в предыдущих периодах позволило оценить долгосрочное влияние инвестиций, а также именно прирост валового регионального продукта как характеристику экономического роста.

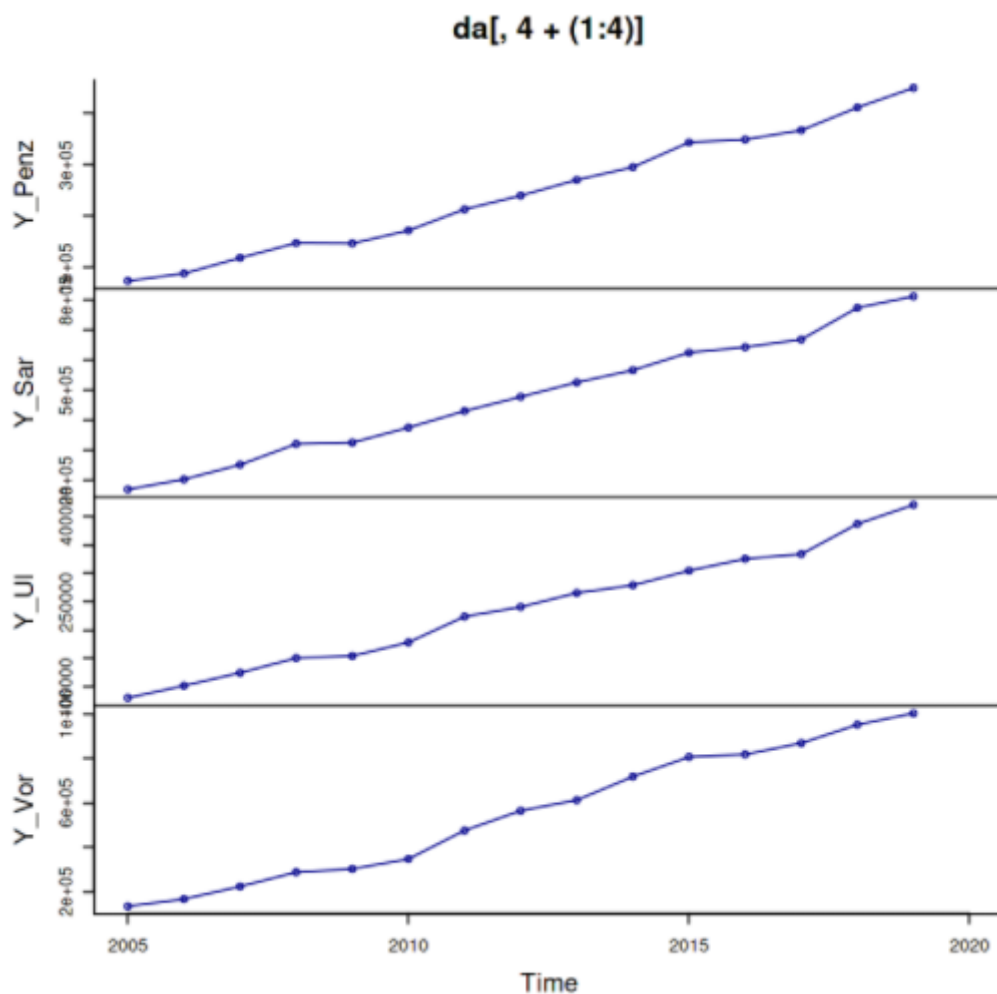
При спецификации эконометрических моделей также учитывалось, что данные как о валовом региональном продукте, так и об инвестициях в основной капитал публикуются как в текущих ценах, так и в виде индексов физического объема. Оценка изменения именно физического объема экономических процессов важна в условиях высокой инфляции, которая в России в разные годы изучаемого временного промежутка составляла от 2,5% (в 2017 году) до 13,3% (в 2008 году). Высокие уровни инфляции делают не вполне корректными выводы по эконометрическим моделям, построенным с использованием показателей в текущих ценах, так как покупательская способность российского рубля за 2005–2019 году существенно изменилась.

В этой связи нами рассчитывались модели с распределенным лагом как для данных о валовом региональном продукте и инвестициях в основной капитал в текущих ценах, так и в сопоставимых ценах 2019 года. Выбор 2019 года в качестве базы продиктован его максимальной близостью к текущему моменту и тем самым наибольшей актуальностью используемых в расчетах уровней цен. Пересчет уровней отдельных лет в сопоставимые цены 2019 года проводился путем деления данных о ВРП последующего года в ценах 2019 года на индекс физического объема.

Во второй главе осуществлены программная реализация и рассчитаны конкретные регрессионные модели ВРП Саратовской, Воронежской, Пензенской и Ульяновской областей. Данные территории выбраны исходя из схожести их территориального расположения в центральной европейской части России, ресурсного потенциала в части полезных ископаемых и человеческого

фактора, экономической специализации на сельском хозяйстве и промышленном производстве.

На основе визуального анализа было сделано предположение о линейном характере изменения ВРП и инвестиций в основной капитал Пензенской (Y\_Penz), Саратовской (Y\_Sar), Ульяновской (Y\_Ul) и Воронежской (Y\_Vor) областей (рис. 1 – пример для ВРП в текущих ценах).



```
plot(da[, 4 + (1:4)], bty="L", type="o", col = "darkblue")
```

Рис. 1. Динамика ВРП Пензенской, Саратовской, Ульяновской и Воронежской областей в текущих ценах, млн. руб.

Алгоритм построения моделей авторегрессии с распределенным лагом включал:

1) построение модели методом наименьших квадратов. На первом этапе в модель включались в качестве результативной переменной ВРП текущего года,

а в качестве объясняющих факторов – ВРП предшествующего года, объем инвестиций в основной капитал текущего и предшествующего годов;

2) проверка значимости уравнения регрессии в целом и каждого из коэффициентов регрессии, визуализация результатов построения модели;

3) в случае отклонения гипотезы о незначимости всего уравнения ( $\alpha=0,05$ ) и незначимости какого-либо фактора за исключением свободного члена – удаление из модели наиболее незначимого фактора;

4) повтор пунктов 1–3 до тех пор, пока все входящие в модель факторы не будут значимы;

5) проверка гипотезы об отсутствии гетероскедастичности на основе теста Вайта;

6) проверка гипотезы об отсутствии автокорреляции остатков на основе тестов Дарбина-Вотсона и Бройша-Годфри.

Пример программой реализации построения и диагностики модели регрессии для Пензенской области представлен ниже.

```
suppressMessages(library(dynlm))
```

```
mdiag <- function(m, bgord=3) {  
  # сводка и диагностика модели регрессии  
  # m модель линейной регрессии  
  # bgord максимальный порядок теста Бройша-Годфри  
  
  # сводка модели  
  print(summary(m))  
  
  # графики  
  options(repr.plot.width = 10, repr.plot.height = 6)  
  par(mfrow=c(2, 2))  
  plot(m)  
  par(mfrow=c(1,1))  
}
```

```

mdiagts <- function(m) {
  # диагностика модели регрессии временного ряда
  # график остатков, тест автокорреляций остатков
  # m модель линейной регрессии

  cat("==== Гомоскедастичность случайных возмущений ====")
  # https://www.r-bloggers.com/2016/01/how-to-detect-heteroscedasticity-and-rectify-it/
  print(lmtest::bptest(m)) # Breusch-Pagan test

  cat("==== Тест Дарбина-Уотсона =====\n")
  print(car::durbinWatsonTest(m))
  cat("\n")

  cat("==== Тест Бройша-Годфри =====\n")
  # https://www.rdocumentation.org/packages/lmtest/versions/0.9-38/topics/bgtest
  print(lmtest::bgtest(m), order=bgord)

  # графики
  options(repr.plot.width = 12, repr.plot.height = 3)
  par(mfrow=c(1, 3))
  plot(m$resid, col="darkblue", bty="L"); abline(h=0, lty = 2)
  acf(m$resid); pacf(m$resid)
  par(mfrow=c(1, 1))
}

```

```

m_eq <- function(m, coe=FALSE, ndig=4) {
  # представляет модель как уравнение LaTeX
  # m модель
  # coe индикатор показа оценочных значений коэффициентов
  # ndig число десятичных знаков после запятой в представлении коэффициентов
  cat("Модель:", deparse(substitute(m)), ifelse(coe, "(оценка)", ""), "\n")
  oldw <- getOption("warn")
  options(warn = -1)
  cat("$$", equatomatic::extract_eq(m, use_coefs = coe, intercept="beta",
                                   wrap = TRUE, coef_digits=ndig), "$$", sep="")
  options(warn = oldw)
}

# Пензенская область, номинальный ВРП, AR(1) + I_Penz + L.I_Penz
m <- dynlm(Y_Penz ~ L(Y_Penz, 1) + I_Penz + L(I_Penz, 1), data=da)
mdiag(m); mdiagts(m); m_eq(m, T)

```

В результате построения регрессионных зависимостей ВРП от объема инвестиций в основной капитал были получены значимые уравнения:

для ВРП в текущих ценах:

Пензенская область

$$Y_{Penz} = -2214.6964 + 0.9828(L(Y_{Penz}, 1)) + 1.1984(I_{Penz}) - 0.7514(L(I_{Penz}, 1)) + \varepsilon; \quad (1)$$

где  $Y_{Penz}$  – ВРП Пензенской области в текущих ценах;



$I\_Penz$  – инвестиций в основной капитал Пензенской области в текущих ценах;

$L(Y\_Penz, 1)$  – лаговое значение ВРП Пензенской области в текущих ценах, лаг =1.

Саратовская область

$$Y\_Sar=13967.7807+0.7596(L(Y\_Sar, 1))+1.2658(I\_Sar) +\varepsilon; \quad (2)$$

где  $Y\_Sar$  – ВРП Саратовской области в текущих ценах;

$I\_Sar$  – инвестиций в основной капитал Саратовской области в текущих ценах;

$L(Y\_Sar, 1)$  – лаговое значение ВРП Саратовской области в текущих ценах, лаг =1.

Ульяновская область – уравнение не удовлетворяет требованиям МНК;

Воронежская область

$$Y\_Vor=30403.5076+0.6299(L(Y\_Vor, 1))+1.2012(I\_Vor) +\varepsilon; \quad (3)$$

где  $Y\_Vor$  – ВРП Воронежской области в текущих ценах;

$I\_Vor$  – инвестиций в основной капитал Воронежской области в текущих ценах;

$L(Y\_Vor, 1)$  – лаговое значение ВРП Воронежской области в текущих ценах, лаг =1.

для ВРП в ценах 2019 года:

Пензенская область

$$Y\_Penz(p2019)=10018.3+0.9018(L(Y\_Penz(p2019), 1))+ \\ +1.0403(I\_Penz(p2019))-0.5889(L(I\_Penz(p2019), 1))+\varepsilon; \quad (4)$$

где  $Y\_Penz(p2019)$  – ВРП Пензенской области в ценах 2019 г.;

$I\_Penz(p2019)$  – инвестиций в основной капитал Пензенской области в ценах 2019 г.;

$L(Y\_Penz, 1)$  – лаговое значение ВРП Пензенской области в ценах 2019 г., лаг =1;

$L(I\_Penz, 1)$  – лаговое значение инвестиций в основной капитал Пензенской области в ценах 2019 г., лаг =1.

Саратовская область

$$Y\_Sar(p2019)=76249.2+0.8625(L(Y\_Sar(p2019), 1))+ \\ +1.0172(I\_Sar(p2019))-0.7731(L(I\_Sar(p2019), 1))+\varepsilon; \quad (5)$$

где  $Y\_Sar(p2019)$  – ВРП Саратовской области в ценах 2019 г.;

$I\_Sar(p2019)$  – инвестиций в основной капитал Саратовской области в ценах 2019 г.;

$L(Y\_Sar(p2019), 1)$  – лаговое значение ВРП Саратовской области в ценах 2019 г., лаг =1;

$L(I\_Sar(p2019), 1)$  – лаговое значение инвестиций в основной капитал Саратовской области в ценах 2019 г., лаг =1.

Ульяновская область – уравнение не удовлетворяет требованиям МНК;

Воронежская область – уравнение не удовлетворяет требованиям МНК.

Уравнение (1) для Пензенской области после переноса в левую часть лаговой результативной переменной будет иметь вид:

$$Y_{\text{Penz}} - 0.9828(L(Y_{\text{Penz}}, 1)) = -2214.6964 + 1.1984(I_{\text{Penz}}) - 0.7514(L(I_{\text{Penz}}, 1)) + \varepsilon; \quad (6)$$

Левая часть уравнения интерпретируется как прирост ВРП к уровню предшествующего года, так как коэффициент 0,9829 при лаговом ВРП близок к 1,0 или к 100%. Коэффициенты при текущем и лаговом объеме инвестиций в основной капитал, расположенные в правой части уравнения  $(1.1984(I_{\text{Penz}}) - 0.7514(L(I_{\text{Penz}}, 1)))$ , могут быть интерпретированы как прирост инвестиций к уровню предшествующего года. Причем за базу сравнения взяты 75,14% инвестиций предшествующего года, так как коэффициент при инвестициях с лагом 1 составляет 0,7514. Для экономической интерпретации целесообразно пересчитать базу сравнения к традиционному 100% уровню.

В целом уравнения (1), (4) и (5) показывают среднее увеличение прироста ВРП в текущих или сопоставимых ценах 2019 года в результате увеличения прироста инвестиций в основной капитал на 1 денежную единицу.

Уравнения (2) и (3) показывают среднее увеличение прироста ВРП в текущих ценах в результате увеличения объема инвестиций в основной капитал текущего года на 1 рубль.

Для данных по Ульяновской области в текущих ценах и по Ульяновской и Воронежской областям в ценах 2019 г. уравнения не были построены, так как коэффициент при лаговом значении ВРП оказался незначим. В этом случае модели показывали бы зависимость не между приростами ВРП и инвестициями, а между исходными показателями. Однако корреляция и регрессия на данных, содержащих тенденцию, является завышенной (ложной) и поэтому исходные значения показателей не могут быть использованы для моделирования.

Также были построены модели с лагами Алмон. Использовалась модель зависимости коэффициентов  $\beta$  от величины лага  $j$  в форме полинома 2-й степени. Максимальная величина лага  $l$  составила три года. Однако ни для

одной из рассматриваемых областей не было построено модели со значимыми коэффициентами регрессии при переменных  $z_1$  или  $z_2$ .

Пример значимой модели Алмон со всеми значимыми коэффициентами для Саратовской области:

$$Y_{Sar} = 28000,9 + 1.19685z_0 + \varepsilon_t \quad \text{или} \quad (6)$$

$$Y_{Sar} = 28000,9 + 1.19685(I_{Sar} + L(I_{Sar}, 1) + L(I_{Sar}, 2) + L(I_{Sar}, 3)) + \varepsilon_t \quad (7)$$

Однако данная модель не может быть использована из-за наличия в исходных данных тенденции и, как следствие, ложной зависимости, а также из-за гетероскедастичности в остатках.

Наличие в исходных данных о ВРП и инвестициях в основной капитал тенденций, схожих с линейной, позволяет строить модели динамических рядов на первых разностях  $d$  ( $d_t = y_t - y_{t-1}$ ). Это позволяет избавиться от ложной связи в результате исключения детерминированного тренда.

Для всех областей как в текущих, так и в сопоставимых ценах 2019 года были получены значимые уравнения регрессии, отличающиеся по величине предсказанной дисперсии результативного показателя. В целом уравнения на первых разностях для показателей в сопоставимых ценах 2019 года имеют большие значения коэффициента детерминации  $R^2$ , чем модели на показателях в текущих ценах.

Динамические модели на первых разностях имеют вид:

Пензенская область

$$dY_{Penz}(p2019) = 11\,171,9 + 0,872124 dI_{Penz}(p2019) + \varepsilon, \quad (8)$$

Саратовская область

$$dY_{Sar}(p2019) = 16\,619,1 + 1,05598 dI_{Sar}(p2019) + \varepsilon, \quad (9)$$

Ульяновская область

$$dY_{Ul}(p2019) = 6\,416,14 + 0,78152 dI_{Ul}(p2019) + \varepsilon. \quad (10)$$

Коэффициент детерминации для уравнения 10 составил 0,31. Для Воронежской области не удалось построить статистически значимого уравнения регрессии. Таким образом, результаты построения моделей динамических рядов схожи с результатами построения моделей автокорреляции с распределенными лагами, с использованием которых также не удалось объяснить экономического роста в Ульяновской и Воронежской областях за счет роста инвестиций.

Коэффициенты регрессии показывают среднее увеличение прироста ВРП в сопоставимых ценах 2019 года (млн. руб.) при увеличении прироста инвестиций в основной капитал на 1 млн. руб.

### **Заключение**

Для осуществления корректного моделирования экономического роста регионов рассмотрены современные методы и инструменты регрессионного анализа, в частности авторегрессионные модели с распределенными лагами, лаги Алмон, динамические модели по первым разностям.

Осуществлен выбор показателей, характеризующих экономический рост регионов и объем инвестиций. В качестве индикатора экономического роста обосновано использование валового регионального продукта, как наиболее обобщающей характеристики развития региона. Для характеристики инвестиций выбраны инвестиции в основной капитал, так как они составляют наибольшую долю от общего объема инвестиций в нефинансовые.

Для исключения инфляции как ВРП, так и инвестиции в основной капитал оценивались не только в текущих ценах, но и в сопоставимых ценах 2019 года. Для пересчета использовались индексы физического объема ВРП и инвестиций в основной капитал, публикуемые Росстатом. Для анализа были собраны данные за 2005–2019 годы.

Осуществлена программная реализация алгоритма эконометрического моделирования зависимости экономического роста от инвестиций на языке R.

Проведен вычислительный эксперимент на примере Саратовской, Пензенской, Воронежской и Ульяновской областей, в результате которого

получены статистически значимые и экономически объяснимые модели. На основе моделей можно говорить, что в Пензенской и Саратовской областях инвестиции используются более эффективно, так как увеличение прироста инвестиций или объема инвестиций приводит к увеличению прироста ВРП. По данным для Ульяновской области в текущих ценах и для Воронежской области как в текущих, так и в сопоставимых ценах 2019 г. не удалось построить значимых моделей, описывающих зависимость между экономическим ростом и инвестициями.

Из-за наличия в исходных данных тенденции в анализе не использовались модели, содержащие только исходные данные о ВРП и инвестициях (текущих или лаговых) для исключения ложной корреляции и, связанных с ней, некорректных выводов. Статистически значимые модели были построены методами авторегрессии с распределенными лагами и МНК на первых разностях. Построение моделей Алмон не дало удовлетворительных результатов.