

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра теории функций и стохастического анализа

**МОДЕЛИ ПРОСТРАНСТВЕННОГО ЛАГА И РЫНОК ТРУДА**

**АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ**

студента 4 курса 412 группы

направления 01.03.02 — Прикладная математика и информатика

механико-математического факультета

Девицына Андрея Алексеевича

Научный руководитель

ст. преподаватель

\_\_\_\_\_

А. Д. Луньков

Заведующий кафедрой

д. ф.-м. н., доцент

\_\_\_\_\_

С. П. Сидоров

Саратов 2022

## ВВЕДЕНИЕ

### **Актуальность темы.**

В настоящее время изучение проблем региональных рынков труда необходимо с точки зрения пространственного взаимодействия. Это особенно важно в условиях России с ее существенными внутрирегиональными различиями, которые влияют на проявления общих закономерностей развития этого рынка и процессов межрегиональных экономических связей.

Межрегиональные связи России и ее субъектов активно оказывают влияние на развитие рыночных отношений через создание конкурентной среды и мотивации для предпринимателей. Уровень развития этих связей страны определяется возможностью хозяйства ее субъектов производить конкурентоспособные на региональном рынке товары и услуги на основе имеющихся конкурентных преимуществ, инвестиционной привлекательности регионов и прочих выгодных для бизнеса возможностей. В регрессионном анализе методы пространственной эконометрики с использованием панельного типа данных позволяют учитывать межрегиональную или пространственную зависимость .

Совокупность панельных данных содержит наблюдения для одних и тех же выборочных единиц. С помощью панельных данных становится возможным специфическое изучение различных социальных проблем таких как безработица, преступность и прочих, полагаясь не только на пространственную выборку или на один временной ряд.

Существует необходимость в учете того, как пространственные эконометрические методы могут определить в совокупности схожесть и различие единиц наблюдения. Зачастую единицы наблюдения, такие как регионы могут обладать пространственным индивидуальным эффектом, который придает уникальность данным, что необходимо учитывать при моделировании и оценивании параметров модели.

В пространственной эконометрике существуют методологические нерешенные вопросы. Одним из таких вопросов является неопределенность в выборе спецификации пространственной взвешивающей матрицы, отражающей взаимосвязи между регионами, а также влияние выбора матрицы на смеще-

ние оценок коэффициентов, который поднимается в данной работе.

**Целью бакалаврской работы** является анализ региональной безработицы России с помощью модели пространственной эконометрики на основе реальных данных.

**Объектом исследования** является региональные рынки труда в России.

**Предметом исследования** являются модели пространственного эконометрического моделирования для анализа и прогнозирования применительно к уровню безработицы в регионах России. Также уделяется внимание проблемам выбора вида пространственного эффекта и весовой матрицы.

Для реализации данной цели были поставлены и последовательно решены следующие **задачи**:

- рассмотрение понятий, моделей и тестов в панельном анализе;
- рассмотрение основных положений в пространственной эконометрике;
- анализ показателей в российском рынке труда;
- выбор и обоснование показателей, как параметров для построения модели;
- построение модели пространственного лага;
- сравнение оценок параметров модели лага.

**Практическая значимость** работы состоит в анализе региональной безработицы по российским данным для пространственной эконометрической модели с обоснованием в выборе спецификации модели и её набора параметров. По результатам этой работы делаются выводы о наличии пространственной зависимости регионов и её влиянии на безработицу.

**Структура и содержание бакалаврской работы.** Работа состоит из введения, четырех разделов, списка использованных источников из 20 наименований, приложений.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснована актуальность темы работы, сформулированы цели и задачи исследования, отмечается практическая значимость полученных результатов работы.

В **первом** разделе приводятся основные понятия связанные с панельными данными и модели со случайными эффектами. Панельные данные состоят из наблюдений одних и тех же экономических единиц или объектов (индивидуумы, домашние хозяйства, фирмы, регионы, страны и т. п.), которые осуществляются в последовательные периоды времени. Они сочетают в себе как данные пространственного типа, так и данные типа временных рядов: в каждый момент времени имеются данные пространственного типа по экономическим единицам, и для каждого такого объекта соответствующие ему данные образуют один или несколько временных рядов.

Панельные данные позволяют выделить характерные различия между индивидуальными экономическими единицами, что выражается в следующей модели:

$$y_{it} = \alpha_i + x_{it}\beta + \epsilon_{it}$$

$y_{it}, i = 1, n, t = \overline{1, T}$ , где  $n \cdot T$  - объем выборки

В модели со случайным эффектом (или random effects model), предполагается, что в этом выражении величина  $\alpha_i = \mu + u_i$ , где  $\mu$  - константа во все моменты времени и  $u_i$  - случайные ошибки, некоррелированные с  $\epsilon_{it}$  и некоррелированы при разных  $t$ .

Благодаря специальной структуре панельные данные позволяют строить более гибкие и содержательные модели и получать ответы на вопросы, которые недоступны только в рамках, например, моделей, основанных на пространственных данных. Также встает вопрос о выборе моделей между фиксированными и случайными эффектами. Модель со случайным эффектом позволяет сделать вывод относительно характеристик генеральной совокупности в отличие фиксированных эффектов, которые не могут рассматривать наблюдения как случайные извлечения из некоторой генеральной совокупности.

Причиной выбора в пользу модели с фиксированным эффектом, это эффекты  $\alpha_i$  при малом числе индивидуальных единиц и при известной природе этих эффектов, а выбор в пользу случайных эффект необходим, когда точно можно сказать об отсутствии корреляции между  $\alpha_i$  и  $x_{it}$  в большой выборке.

Чтобы внести ясность в этот выбор нужны статистические тесты для проверки гипотез. Тест Хаусмана сравнивает состоятельные оценки нулевых и альтернативных гипотез и при существенной разнице в значениях будет свидетельствовать о справедливости нулевой гипотезы. Тест Бреуша-Пагана используется для проверки гетероскедастичности, а  $F$ -тест используется для модели с фиксированными эффектом с нулевой гипотезой  $H_0 : \alpha_1 = \dots = \alpha_n$ .

Случайные эффекты можно рассматривать как модель с фиксированными в случае отсутствия корреляции между эффектами и регрессорами или рассматривать как частный случай обыкновенной регрессионной модели.

В **втором** разделе рассмотрены основные положения пространственной эконометрики, а именно принцип построения весовых матриц, пространственной модели с лагом и оценка этой модели методом максимального правдоподобия.

В анализе индивидуальных экономических объектов, территориально привязанных к своему местоположению, неизбежно взаимное пространственное влияния, игнорирование которого приведет к некорректным выводам в отношении значимости самого влияния.

Весовая матрица количественно описывает пространственные и временные отношения, которые существуют между объектами в наборе данных.

В простейшем случае элементы матрицы будут принимать только два значения, зависящие от наличия связи между объектами, то непосредственно территориальное соседство:

1.  $w_{ij} = 1$ , если  $i$  граничит с  $j$ ;
2.  $w_{ij} = 0$ , если  $i$  не граничит с  $j$  или исключено влияние единицы на себя  $i = j$ .

Такая матрица называется матрицей соседства. Однако в этом предполагается, что на экономику объекта влияют только граничащие с ней единицы. Некоторой модификацией этой матрицы является матрица  $k$  ближайший со-

седей, для построения матрицы вычисляется значение расстояния от одной единицы до других и выбираются несколько наименьших значений в качестве  $k$  ближайших соседей.

Как альтернатива матрицы соседства существует матрица обратных расстояний между объектами, которая имеет следующие условия:

1.  $w_{ij} = 0$ , если  $i = j$  или  $d_{ij} > D(q)$ ;
2.  $w_{ij} = \frac{1}{d_{ij}^\lambda}$ , если  $D(q) \leq d_{ij}$ .

Необходимо отметить, что необходимость интерпретации пространственного влияния ведет к построению взвешивающей матрицы и любой подход к её формированию сводится к комбинации учета факторов влияния, которые нужно обосновать.

При построении обыкновенной регрессионной модели рассматривают матрицу независимых переменных  $X$ , эта модель не учитывает индивидуальный эффект единиц наблюдения. Так как наблюдение должно представлять из себя регион(область), то матрице  $X$  необходима параметр пространственной зависимости. Для этого рассматривается модель с пространственным лагом(Spatial autoregressive model или SAR)

$$y_{it} = \rho \sum_{j=1}^n W_{ij} y_{jt} + x_i \beta + \epsilon_{it}$$

где  $\sum_{j=1}^n W_{ij} y_{jt}$  — пространственный лаг, представляющий собой линейную комбинацию значения показателя  $y_j$  — соседних наблюдений  $i$  с элементами матрицы пространственных весов  $W_{ij}$ .

Коэффициент  $\rho$  для пространственного авторегрессионного процесса — это коэффициент корреляции между зависимой переменной и пространственным лагом. Если коэффициент корреляции положительный и существенно отличается от нуля, то существует положительная автокорреляция. Если коэффициент близок к нулю, то автокорреляция будет отсутствовать.

**Третий** раздел посвящен анализу показателей и факторов влияющих на рынок труда в России. Современное социально-экономическое развитие регионов Российской Федерации характеризуется крайней неравномерностью и зависит от множества факторов, определяющих потенциальные возможно-

сти территорий. Для регионов страны характерно неравномерное размещение различных видов ресурсов, накопленных элементов национального богатства, объектов инфраструктурного обеспечения. По данным доклада Всемирного банка от 2019 года, Россия находится на третьем месте в рейтинге стран Европы и Центральной Азии по степени неравенства регионов, которое проявляется на всех уровнях – от федеральных округов до муниципальных образований. Это оказывает влияние на формирование территориального разделения труда, в соответствии с которым каждый регион может предложить на рынок те виды продукции и услуг, для производства которых он имеет наиболее благоприятные условия.

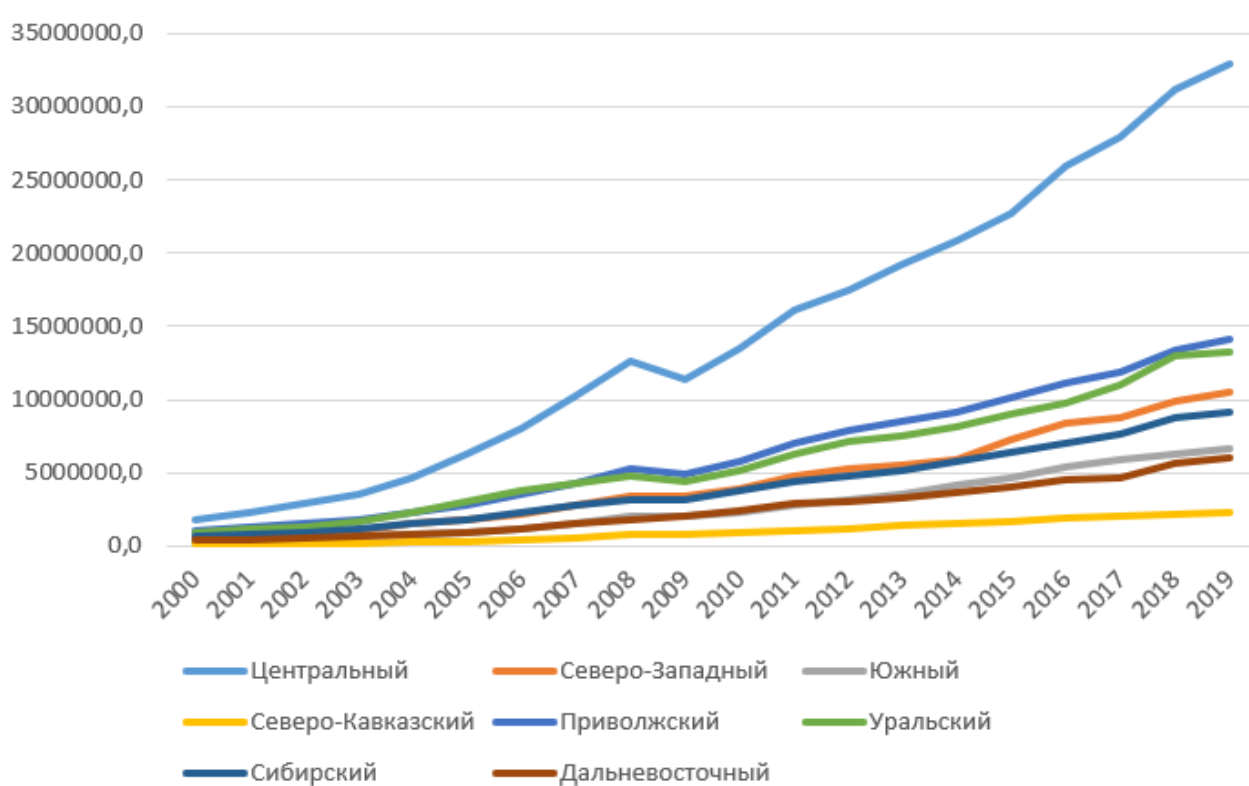


Рисунок 1 – ВРП по федеральным округам

Как можно заметить на рисунке 1 разница между Центральным и остальными округами по показателю ВРП существенная, однако тенденция к росту ВРП по округам одинакова. Данная иллюстрация только одно из следствий регионального неравенства, следовательно его можно рассматривать как фактор влияния на уровень безработицы в округах.

Посмотрев на российский рынок труда в разрезе численность рабочей

силы, то можно заметить, что в России с 2016 года начинает заметно снижаться численность рабочей силы. За последние четыре года с 2016 по 2020 годы численность рабочей силы снизилась с 76 636 тыс. чел. до 74 923 тыс. чел., и продолжает снижаться дальше.

Таблица 1 – Численность в тыс. человек

	2015	2016	2017	2018	2019	2020
Численность рабочей силы	76588	76636	76285	76190	75398	74923
Численность безработных	4264	4243	3969	3658	3465	4321

Среди причин можно выделить эмиграцию работников, выход на пенсию и последствия пандемии. На фоне этого флуктуация уровня безработицы среди трудоспособного возраста находилась в пределах 1%.

Таблица 2 – Уровень безработицы(в процентах)

	2015	2016	2017	2018	2019	2020
всего	5.6	5.5	5.2	4.8	4.6	5.8
в трудоспособном возрасте	5.82	5.78	5.39	5.02	4.85	6.10
зарегистрированная безработица	1.3	1.2	1	0.9	0.9	3.7

В ситуации снижения численности трудовых ресурсов возможно, что работодатели будут все в большей мере сталкиваться с нехваткой рабочей силы.

В **четвертом разделе** идет обоснование выбора весовой матрицы и оценка параметров модели пространственного лага с дальнейшим ее построением по данным Федеральной службы государственной статистики.

Выбор взвешивающей матрицы должен быть обоснован, но в экономической теории нет однозначного ответа на этот вопрос и даже не существует какого-либо эмпирического теста для его нахождения. Поэтому способом проверить наличие или отсутствие какой-либо корреляции между наблюдениям (пространственного взаимовлияния между регионами) возможно с помощью непосредственного сравнения оценок одной и той же модели с использованием разных весовых матриц.

В российской статистике существует два потенциальных источника данных о рынке труда. Первый из них это данные предприятий, собираемой



Федеральная служба государственной статистики. Второй — это данные, заявленные предприятиями в государственную службу по труду и занятости в целях поиска персонала. Оба источника имеют свои недостатки. В первом случае покрываются только крупные и средние предприятия, и при этом данные могут содержать много номинальных вакансий, не имеющих отношения к реальному спросу на труд. Во втором случае статистика покрывает только те предприятия, которые рассматривали службу по труду и занятости как потенциальный способ поиска работников, и при этом список вакансий обычно смещен в сторону представителей рабочих специальностей, которые заняты в промышленной и производственной сфере.

Одним из факторов, учтенных в модели является отраслевая структура экономики. Следующим является уровень образования. Люди с высоким уровнем образования имеют больше шансов быть принятыми на работу в других регионах, вследствие чего чаще мигрируют, чем люди с низким уровнем образования. Вследствие этого при большой доле людей с высшим образованием в регионе будет снижаться уровень безработицы. Другим существенным фактором для уровня безработицы является рабочая сила и она, как ранее было замечено в России, снижается в абсолютном значении, что приводит к увеличению безработицы.

### Вычислительный эксперимент

В данной работе **целью вычислительного эксперимента** является оценка параметров модели пространственного лага (SAR) с разными взвешивающими матрицами на основе реальных данных. Данные были взяты с портала Федеральной службы государственной статистики с 2015 по 2019 год для 75 российских регионов (исключением являются некоторые автономные округа), построение модели реализовано в статистическом программном пакете STATA.

Для реализации цели используется модель пространственного лага (SAR) следующего вида:

$$y_{it} = \rho \sum_{j=1}^n W_{ij} y_{jt} + x_i \beta + \epsilon_{it} + \mu_i$$

где  $i = \overline{1, n}$ , номер региона, а всего регионов  $n = 75$ ,  $t = 2015 \dots, 2019$  года,  $\mu_i$  – индивидуальный эффект,  $x_i$  – объясняющие переменные,  $W$  – взвешивающая матрица размерности  $n \times n$ , а ошибки  $\epsilon_{it} \sim (0, \sigma^2)$ . Список объясняющих переменных для модели представлен в таблице 3:

Таблица 3 – Описание независимых переменных

Описание переменной	Обозначение переменной
Доля отрасли «обрабатывающие производства» в валовой добавленной стоимости	man_ind
Доля отрасли «добыча полезных ископаемых» в валовой добавленной стоимости	min_ext
Доля государственных услуг (образование, здравоохранение, государственное управление) в валовой добавленной стоимости	ser_gov
Доля негосударственных услуг (гостиницы, транспортные услуги, финансовая деятельность, операции с недвижимостью) в валовой добавленной стоимости	non_gov
Доля занятых с высшим образованием	epm_high
Доля населения в трудоспособном возрасте, от общего количества населения	work_pop
ВРП на душу населения в ценах базового года	vtp_soul

Федеральная служба государственной статистики описывает валовую добавленную стоимость «на уровне отраслей и секторов как разность между выпуском товаров и услуг и промежуточным потреблением», что отражает

различие в спросе на труд в долях между регионами. Однако, указанные доли в таблице 3 не представляют всю структуру валовой добавленной стоимости, чтобы избежать линейную зависимость при расчетах в модели.

Сравниваются коэффициенты при независимых переменных моделей SAR и делается вывод о том, что нет больших различий в оценках коэффициентов и их значимости между моделями одной спецификации с разными взвешивающими матрицами.

Тест Бреуша-Пагана показывает наличие случайного индивидуального эффекта модели, поскольку по результатам теста нулевая гипотеза  $Var(u) = 0$ , отвергается.

Далее используется тест Хаусмана для определения наличия корреляции между независимыми переменными и эффектами для модели с одним типом, а затем с разным типом взвешивающей матрицы. Из результатов теста следует, что стоит сделать выбор в пользу модели SAR с фиксированными эффектами так как состав выбранных регионов не менялся за указанный период, а матрица обратных расстояний учитывает влияние всех регионов друг на друга.

## Основные результаты

1. рассмотрены понятия, моделей и тестов в панельном анализе;
2. рассмотрены основные положения в пространственной эконометрике;
3. проведен анализ показателей в российском рынке труда;
4. проведен выбор и обоснование показателей, как параметров для построения модели;
5. построены модели пространственного лага с параметрами;
6. результаты сравнения оценок и тестов показал не существенную разницу между взвешивающими матрицами.