

Министерство науки и высшего образования Российской Федерации  
федеральное государственное автономное образовательное учреждение  
высшего образования  
«Санкт-Петербургский политехнический университет Петра Великого»  
(ФГАОУ ВО СПбПУ)  
Институт промышленного менеджмента, экономики и торговли  
Высшая инженерно – экономическая школа

## КУРСОВАЯ РАБОТА

По дисциплине «Эконометрика: продвинутый уровень»

Эконометрический анализ и моделирование интернет-продаж  
в Российской Федерации  
(семестр 2)

Студент группы  
3740105/20101

---

подпись, дата

К.С. Малышева  

---

инициалы и фамилия

Оценка выполненной студентом работы:

Преподаватель,  
доцент, к.э.н.

---

подпись, дата

А.Е. Схведиани  

---

инициалы и фамилия

Санкт-Петербург – 2023

САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ  
ПЕТРА ВЕЛИКОГО

Институт промышленного менеджмента, экономики и торговли  
Высшая инженерно – экономическая школа

**ЗАДАНИЕ**

на выполнение курсовой работы по дисциплине «Эконометрика:  
продвинутый уровень»

студенту Малышевой К.С.

группа: 3740105/20101

семестр: 2

1. Тема работы: Эконометрический анализ и моделирование интернет-продаж в Российской Федерации

2. Срок сдачи студентом законченной работы: 25.05.2023

3. Исходные данные по работе: 1) Теоретические материалы лекций. 2) Учебная литература и информация, представленная на сайтах. 3) Программное обеспечение Python и другие программные продукты по выбору студента. 4) Статистические данные портала <https://coinmarketcap.com/>. 5) Статистические данные по регионам Единой межведомственной информационно – статистической системы (ЕМИСС): <https://www.fedstat.ru/>. 6) Статистические данные из базы данных World bank: <https://data.worldbank.org/>. 7) Статистические данные из базы данных OECD data: <https://data.oecd.org/>. 8) Статистические данные Федеральной службы государственной статистики (Росстат): <https://rosstat.gov.ru/folder/210/document/13204>. 9) База данных по Российским компаниям - СПАРК 10) ИНИД: <https://www.data-in.ru/about/>. 11) Kaggle: <https://www.kaggle.com/>. 12) Harvard Dataverse: <https://dataverse.harvard.edu/>. 13) Google data search: <https://datasetsearch.research.google.com>.

4. Содержание работы (перечень подлежащих разработке вопросов): 1. Анализ литературных источников по теме исследования. 2. Описание методов и данных, используемых в исследовании. 3. Представление результатов эконометрического анализа. 3.1. Анализ описательной статистики: выявление закономерностей и особенностей выборки. 3.2. Анализ результатов эконометрического моделирования: сравнение моделей, оценка значимости коэффициентов, обоснование выбора наилучшей модели. 4. Обсуждение результатов: сравнение

полученных результатов с результатами других авторов, представленных в анализе литературных источников.

5. Перечень графического материала (с указанием обязательных рисунков): 1. Результаты систематизации факторов и их влияния на зависимую переменную согласно анализу литературных источников. 2. Описание переменных: обозначение, краткое описание, единицы измерения. 3. Результаты анализа описательной статистики 4. Графики распределения значений признаков относительно исследуемых объектов. 5. Корреляционная матрица, содержащая оценки коэффициентов корреляции. 6. Сравнение результатов регрессионного анализа по различным моделям. 7. Результаты моделирования линейно – логарифмической, логарифм-линейной и логарифмической моделей. 8. Результаты проверки на мультиколлинеарность с помощью корреляционной таблицы. 9. Результаты проверки на мультиколлинеарность с помощью показателя VIF. 10. Результаты проверки на линейность с помощью частных остаточных графиков. 11. Результаты проверки на гетероскедастичность с помощью графика рассеивания остатки-предсказанные значения. 12. Результаты проверки на гетероскедастичность с помощью графика рассеивания предсказанных и исходных значений  $Y$ . 13. Результаты проверок на основные предпосылки регрессионного моделирования с использованием статистических тестов.

6. Дата выдачи задания: 07.02.2023

Преподаватель

А.Е. Схведиани

\_\_\_\_\_  
(подпись)

\_\_\_\_\_  
(инициалы, фамилия)

Студент

К.С. Малышева

\_\_\_\_\_  
(подпись)

\_\_\_\_\_  
(инициалы, фамилия)

## СОДЕРЖАНИЕ

АННОТАЦИЯ ИССЛЕДОВАНИЯ .....	5
ВВЕДЕНИЕ .....	6
ОСНОВНАЯ ЧАСТЬ.....	8
1. Литературный обзор .....	8
2. Данные и методы.....	23
2.1. Описание исходных данных .....	23
2.2. Методы и модели исследования.....	23
3. Результаты анализа .....	25
3.1. Результаты анализа описательной статистики .....	25
3.2. Результаты эконометрического анализа.....	30
4. Обсуждение результатов .....	33
ЗАКЛЮЧЕНИЕ .....	36
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ .....	37
ПРИЛОЖЕНИЕ А. Исходные данные исследования.....	40
ПРИЛОЖЕНИЕ Б. Код исследования.....	41
ПРИЛОЖЕНИЕ В. Промежуточные результаты расчетов и проверок на гипотезы .....	47

## **АННОТАЦИЯ ИССЛЕДОВАНИЯ**

Развитие информационных технологий и телекоммуникаций открыло новые возможности для бизнеса. Эти поразительные события привели к быстрому распространению электронной коммерции, которая набирает популярность во всем мире и способствует экономическому росту страны. Целью исследования является эконометрический анализ возможных факторов, способствующих росту или сдерживанию интернет-продаж в Российской Федерации на основе данных из открытых источников. Для этого были получены данные по 81 субъекту РФ за периоды с 2016-2021 гг. Было предложено семь различных моделей множественной регрессии путем сочетания различных факторов, чтобы проанализировать их относительное влияние на рост электронной коммерции. По результатам исследования выдвинутая гипотеза в работе была доказана не полностью. Чувствительность логарифма доли интернет-продаж в РФ была самой высокой для доли молодого населения, логарифма доли компаний и населения, использующих интернет, пандемии коронавируса. Также значимыми переменными оказались логарифм среднедушевых денежных доходов населения, индекса потребительских цен, доли городского населения, логарифм протяженности автомобильных дорог с покрытием, цифровизации телефонной сети в городах, уровня образования населения, доли инвестиций в исследование и развитие технологий.

## **ВВЕДЕНИЕ**

Стремительное развитие цифровых технологий и их массовое внедрение в различные области человеческой жизнедеятельности ознаменовали собой новый этап в развитии социально-экономических процессов общества. Прежде всего, это связано с появлением цифровой экономики, в рамках которой возникли новые формы рыночного взаимодействия, в частности, такая как электронная коммерция [1].

Электронная коммерция – это транзакция, которая совершается в сети компьютеров, связанных между собой, после завершения которой определяются права пользования услугой или товаром [2].

Под электронной торговлей подразумевается деятельность экономических субъектов, связанная с реализацией коммерческих операций с использованием электронных средств обмена информацией [2].

В сравнении с традиционным электронный обмен данными в Интернете является более эффективным, причем как по скорости прохождения, так и по себестоимости. Благодаря ему пользователи могут обмениваться данными, осуществлять финансовые расчеты и заключать сделки в онлайн-режиме [2].

Рынок интернет-торговли является относительно новым явлением для российской экономики, но уже сформировалась устойчивая динамика активного развития данного сектора. Доля товаров, проданных через интернет, динамично увеличивается в суммарном обороте розничной торговли России, что подтверждает исследование Ассоциации компаний интернет-торговли. На это развитие влияет множество показателей как на уровне страны, так и на уровне участников интернет-торговли [3]. Для исследования закономерностей в данной сфере необходимо выделить наиболее значимые факторы, определяющие векторы и темпы развития рынка. Но большинство поздних исследований в области электронной коммерции заинтересованы в определении факторов, специфичных для фирмы и клиента, в то время как ни одно из них не принимало во внимание

национальные факторы, которые могут увеличить деловую активность в Интернете, что подтверждает актуальность рассматриваемой в данной работе тематики.

Целью курсовой работы является моделирование доли розничных интернет-продаж по данным субъектов Российской Федерации и систематизирование факторов, оказывающих воздействие на его темпы роста.

Для достижения обозначенной цели необходимо решить следующие задачи:

- создание модели, описывающей ситуацию на рынке интернет-торговли по данным субъектов Российской Федерации;
- выделение и распределение по степени значимости факторов интернет-продаж по Российской Федерации.

Объектом исследования являются субъекты Российской Федерации.

Предмет исследования – интернет-продажи по субъектам Российской Федерации.

Методология исследования: теоретические (анализ, синтез, обобщение) и эмпирические методы (наблюдение, сравнение, описание), множественный регрессионный анализ

Основными инструментами исследования являются программное обеспечение Excel и Jupyter Notebook.

## ОСНОВНАЯ ЧАСТЬ

### 1. Литературный обзор

Развитие информационных технологий и телекоммуникаций открыло новые возможности для бизнеса. Эти поразительные события привели к быстрому распространению электронной коммерции, которая набирает популярность во всем мире и способствует экономическому росту страны. Рынок интернет-торговли является динамично развивающимся как в России, так и в мире, что подтверждают исследования различных агентств (McKinsey, Remarkety, Data Insight, АКИТ). Причинами этому могут быть различные факторы. В научном сообществе имеют предлагается разный набор показателей, отражающих развития российской экономики в сети Интернет.

Для изучения возможных факторов влияния на рынок необходимо для начала определить общую тенденцию его развития.

Сегодня много говорится о том, что электронная коммерция развивается очень быстрыми темпами, что в ближайшем будущем она вытеснит традиционные формы торговли. Однако конкретных числовых данных о состоянии нового рынка приводится не так и много, при этом официальной информации практически нет, а данные представлены какими-либо аналитическими агентствами [4].

По данным большинства аналитических агентств, на протяжении последних пяти лет оборот розничной Интернет-торговли, рос достаточно динамично (хотя происходило постепенное замедление, обусловленное изменением фаз развития рынка) – в среднем на 30% в год, включая сложный 2015 г. При этом онлайн сегмент существенно опережал российский ритейл. Оборот розничной торговли с 2011 по 2014 гг. рос умеренными темпами – около 11% в год, а в 2015 г. вовсе сократился на более чем 30% по причине тяжелой макроэкономической ситуации в стране [5].

Данную динамику объясняют эффектом «низкой базы», Россия начала свой путь в электронной коммерции позже, чем зарубежные организации, но



основываясь на опыте зарубежных компаний, что при правильном переносе этого опыта на российский рынок позволяет показывать стремительную динамику развития [6].

Российский рынок электронной коммерции демонстрирует рост экспоненциальными темпами. Причем среднегодовые темпы роста даже превышают темпы роста сферы e-commerce в мире. По объему рынка электронной торговли Россия занимает девятое место в первой десятке крупнейших мировых рынков электронной коммерции. Так, среднегодовой цепной темп роста российского рынка составляет 1,2056, а за последние 7 лет рынок вырос более чем в три раза. При этом ожидается, что в 2020 г. он вырастет до уровня 2 трлн руб., хотя по-прежнему останется примерно в районе 0,5% от мирового уровня рынка e-commerce [4].

Таким образом, в среднем доля электронной коммерции ежегодно увеличивается на 0,33%, что явно недостаточно по сравнению со среднемировым показателем (около 0,8%) при более высоком стартовом значении. Пока можно говорить о том, что в России еще не созданы условия для развития электронной коммерции, которые сложились в мире даже до современного мирового уровня доли электронной коммерции – около 10% при сохраняющихся темпах роста около 15 лет [4].

Согласно данным годового отчета, представленного компанией «Statistica», специализирующейся на прогнозах, аналитики рынка и услуг, развитие интернет-торговли в ближайшие пять лет с учетом влияния всех положительных и отрицательных факторов составит 33,2%, в сумме к 2024 г. совокупная выручка от онлайн-продаж предположительно составит 7,2 трлн рублей [7].

По мнению аналитиков из компании McKinsey, данная положительная тенденция сохранится, а «к 2025 году потенциальный экономический эффект от цифровизации экономики России может увеличить ее ВВП на 4,1–8,9 трлн руб. (в ценах 2015 г.), что составит от 19 до 34% общего ожидаемого роста ВВП [8].

Данные исследовательских компаний свидетельствуют о том, что цифровая экономика в России является одним из наиболее активно развивающихся направлений. По разным оценкам, ее доля в ВВП страны выросла с 2,4% в 2015 г. до 2,6% в 2016 г., за аналогичный период времени доля зависимых от цифровой экономики отраслей в ВВП увеличилась с 10 до 16% [8].

Таким образом, можно утверждать, что российский рынок электронной торговли демонстрирует положительную динамику роста. Далее необходимо определить основные факторы, которые оказывают влияние на данную тенденцию.

Различные исследования данной темы с точки зрения эконометрического анализа и построение регрессионной модели можно найти как у российских, так и зарубежных авторов. В основном исследования проводятся на уровне отдельной страны или региона, но все же встречается пересечение между предлагаемыми факторами воздействия на темпы рынка интернет-торговли.

Для начала обратимся к зарубежным статьям и рассмотрим модели, взгляды, предлагаемые иностранными исследователями.

В своей статье [9] группа аналитиков из разных стран мира выдвигают девять различных моделей множественной регрессии потенциала электронной коммерции любой страны. В качестве показателя потенциала электронной коммерции (зависимой переменной в регрессионной модели) авторами был взят индекс электронной торговли В2С ЮНКТАД из Доклада об информационной экономике, который в свою очередь оценивает возможности страны по ведению электронной коммерции. В качестве переменных были исследованы и объяснены три группы факторов:

- группа технологических факторов, которая включает данные инфраструктуры ИКТ, мобильного телефона и виртуальной социальной сети из Глобального отчета по информационным технологиям за 2014 опубликованным Всемирным экономическим форумом;

- группа социальных факторов, которая включает в себя показатели урбанизации (данные о городском населении в процентах от общей численности населения с сайта Всемирного банка), уровень образования (оценка ожидаемого года обучения из Доклада о человеческом развитии) и кибербезопасности (глобальный индекс кибербезопасности (GCI), разработанный ABI Research);

- группа экономических факторов, которая включает показатели уровня доходов (ВНД на душу населения из Доклада о человеческом развитии), логистическое развитие (индекса эффективности логистики) и возможности для бизнеса (оценка «Расстояние до рубежа» из отчета «Ведение бизнеса», опубликованного Всемирным банком).

Все данные по исследуемым показателям были собраны по 145 странам мира. За базовый год для сбора данных был взят 2014 год. Авторами статьи были выдвинуты следующие гипотезы:

1) различные технологические факторы, такие как инфраструктура ИКТ, количество мобильных абонентов и пользователей социальной виртуальной сети, положительно связаны с потенциалом электронной коммерции;

2) социальные факторы, такие как городское население, ожидаемая продолжительность обучения и безопасность от киберпреступлений, положительно связаны с потенциалом электронной коммерции;

3) экономические факторы, такие как ВНД на душу населения, показатель индекса эффективности логистики и показатель расстояния до границы, положительно связаны с потенциалом электронной коммерции;

4) различные технологические факторы (инфраструктура ИКТ, количество мобильных абонентов и пользователей социальных виртуальных сетей), социальные факторы (городское население, ожидаемая продолжительность обучения и безопасность от киберпреступности) и экономические факторы (ВНД на душу населения, логистические Оценка

индекса производительности и оценка расстояния до границы) положительно связаны с потенциалом электронной коммерции.

Исследователями было построено девять моделей множественной регрессии: четыре модели были использованы для проверки гипотез, в то время как остальные – разрабатывались путем комбинации различных переменных.

Наибольшую значимость по коэффициенту  $R^2$  имеет модель, проверяющая четвертую гипотезу: значение составило 0,923. Эта модель показала, что увеличение на единицу ВНД на душу населения и индекса ИКТ приводит к соответственному увеличению стоимости электронной коммерции на 15,6 и 8,6 единицы. В той же модели использование сайтов виртуальных сетей и городское население связаны положительно, тогда как оценка подписки на мобильные телефоны связана отрицательно. Остальные четыре переменные были статистически незначимы. Таким образом, в результате исследования авторы сделали следующие выводы:

- ВНД на душу населения был наиболее важным фактором роста онлайн-бизнеса; вторым по важности фактором был индекс ИКТ (который представляет собой агрегированный показатель уровня инфраструктуры ИКТ, доступности и навыков);

- такие переменные, как городское население и виртуальные сетевые сайты, были важным предиктором потенциала электронной коммерции, однако результаты таких факторов, как показатель эффективности логистики и ожидаемые годы обучения, были неоднозначными, поскольку иногда эти переменные становятся статистически незначимыми.;

- к полному удивлению, кибербезопасность оказала убедительное влияние на электронную коммерцию в немногих случаях, но чаще во многих оно оказывалось несущественным, точно так же влияние деловых перспектив оказалось очень незначительным и не поддающимся обобщению.

В статье [10] группа исследователей из Восточной Европы предлагает регрессионную модель электронной коммерции, основанную на данных

стран Европейского Союза и ее кандидатов. В анализе зависимой переменной являются интернет-покупки физических лиц за предыдущие 12 месяцев 2015 года, приведенные в процентах от общего числа физических лиц в стране. В целом в исследовании была рассмотрена 31 европейская страна: 28 государств - членов Европейского союза, БЮРМ, Сербия и Турция.

В качестве переменных на основе предыдущих исследований авторами были выделены следующие переменные: ВВП на душу населения в процентах (база данных Евростата), государственные расходы на образование (в процентах от ВВП по данным Евростата и Всемирного банка), уровень проникновения Интернета (процентная доля лиц, пользующихся Интернетом) и уровень интернет-навыков отдельных лиц (процентная доля лиц в возрасте от 16 до 74 лет, которые смогли выполнить 1 или 2 из 6 видов деятельности, связанных с Интернетом по данным Евростата).

Исследователями были отобраны и построены четыре регрессионные модели с разным набором переменных, где поэтапно включалось от одной, до четырех переменных. Проанализировав результаты, авторы сделали следующие выводы:

- покупки в Интернете как показатель уровня электронной коммерции сильно коррелируют с использованием Интернета, уровнем ВВП и государственными расходами на образование (уровень значимости всех трех переменных был выше 5%);

- наиболее сильное прямое влияние на электронную коммерцию оказывает уровень проникновения Интернета (использования Интернета) среди населения страны.

Подобное исследование европейского рынка было проведено группой авторов в статье [11]. Авторами были рассмотрены данные по отобранным европейским странам в период с 2003 года по 2017 год. В выборку попали следующие государства: Бельгия, Германия, Ирландия, Греция, Испания, Литва, Австрия, Польша, Португалия, Финляндия, Швеция, Великобритания

и Норвегия. Данные взяты из Eurostat (2019a, 2019b) и WB (2018a, 2018b, 2018c).

Для данного исследования развитие электронной коммерции было выражено через показатель общего товарооборота компаний в сфере интернет-торговли в процентном отношении от общего оборота продаж.

В качестве детерминантов были отобраны следующие показатели:

- количество подписок на мобильный Интернет на 100 человек;
- процент от ВВП, посвященный исследованиям и расходы на развитие (RD) в каждой стране;
- располагаемый доход на душу населения из каждой страны.

В результате исследования была доказана взаимосвязь между анализируемыми показателями и их положительное влияние на объем товарооборота интернет торговли. Наиболее значимым показателем оказался реальный располагаемых доход населения, на втором месте – количество подписок на мобильный Интернет.

Рассмотрим исследование испанского рынка интернет-торговли, проанализированного в работе [12]. Авторы пытаются также выявить и систематизировать возможные факторы, оказывающие влияние на эволюцию электронной коммерции через построение регрессионной модели.

Исследователями анализировались данные по отобранным показателям в период с 1996 г. по 2003 г.

В качестве зависимой переменной авторами был выбран показатель объема продаж в сфере электронной коммерции B2C в Испании, выраженный в миллионах евро. Данные по этому показателю были получены в результате опроса испанского населения старше 14 лет, ежегодно разрабатываемого с 1996 года АЕСЕМ (Испанской ассоциацией электронной коммерции) и Red.es Обсерватория (государственная фирма правительства Испании, которая способствует развитию информационного общества).

В качестве детерминантов были отобраны следующие показатели:

- потенциальный объем рынка, выраженный через количество пользователей Интернета в Испании в тысячах пользователей (данные из отчета Испанской ассоциации исследований средств массовой информации);

- технологический фактор (распространенность широкополосной связи) и правовой фактор (правовые рамки защиты прав потребителей в Интернете), выраженные через дихотомическую переменную, которая равна 0 за период с 1996 по 2001 год и 1 с 2002 года и далее, так как в 2002 году была начата разработка Закона об услугах информационного общества и электронной торговле, регулирующий сферу электронной коммерции, а также о распространении связи;

- среднерыночная стоимость на самый производительный компьютер в евро в Европе (по данным, ежегодно предоставляемые Европейской обсерваторией информационных технологий);

- оценка фактора безопасности в ценностном предложении фирм в Интернете, выраженная через показатель количества серверов с доменом ".es" - кодом страны для Испании - которые используют протоколы безопасности на 1000 пользователей (по данным, подготовленным для Испании сервисной фирмой Netcraft).

Авторами была построена модель множественной регрессии, которая показала положительную зависимость между объемом продаж в Интернете и выбранными детерминантами ( $p$ -value не равно 0), при этом коэффициент  $R^2$  составляет 95,9%. При этом решающее значение имеет показатель потенциального объема рынка.

В заключении авторами было отмечено, что со временем развития интернет-торговли состав определяющих факторов будет претерпевать изменения. Например, большее влияние будет оказывать структура пользователей Интернета (городское население, с университетским образованием, уровнем дохода выше среднего по населению и с компьютерным оборудованием дома).

Анализ испанского рынка интернет-торговли был проведен в работе [13]. Авторы рассматривали влияние социально-демографических факторов на объем электронной коммерции в Испании. Для исследования были взяты данные с помощью опроса, проведенного Национальным статистическим институтом Испании (Instituto Nacional de Estadística, INE), в рамках которого ежегодно собирается информация об ИКТ-оборудовании испанских домохозяйств. Авторами была проанализирована информация с января по май 2014 года; и с февраля по май 2019 г. За 2014 г. выборка состоит из 2487 респондентов (INE, 2014), а на 2019 год — 6308 респондентов (INE, 2019).

В качестве зависимой переменной был выбран фиктивный показатель, который показывает 0 – если респондент не связан с электронной коммерцией и 1 – если является активным пользователем.

В качестве детерминантов были выбраны следующие данные:

- способности к ИКТ (использование компьютера и Интернета): варьируются от низкого уровня до очень высокого;
- демографические показатели: пол, возраст, гражданство;
- социальные показатели: уровень образования, доход;
- уровень безопасности: выражен в чувстве уверенности и в доверии к электронным технологиям, варьируется от низкого до высокого;
- наличие подключения к широкополосному Интернету.

В ходе исследования авторами были сделаны следующие выводы:

- со временем важность социально-демографических показателей таких, как уровень образования, дохода, возраст и пол снижается;
- наиболее значимыми факторами являются навыки пользования ИКТ и доступность подключения к широкополосному Интернету, а также уровень электронного доверия, который вытекает из двух указанных показателей.

Также рассмотрим исследование индийского рынка электронной торговли, представленное в работе [14]. Анализируемой зависимой переменной был выдвинут показатель объем розничных продаж через Интернет в Индии за период с 2005 года по 2020 год. В качестве переменных



были отобраны следующие показатели: количество пользователей Интернетом, количество пользователей кредитными и дебетовыми картами, темпы роста ВВП.

Авторами были использованы модель двойной логарифмической множественной регрессии, тест причинности Грейнджера, тест коинтеграции Йохансена (1988) и модель VEC (1996).

По итогу проведенного исследования в статье делается вывод о том, что однопроцентное увеличение доли населения, пользующегося Интернетом, привело к 1,931% увеличению розничных продаж в электронной коммерции в Индии в 2005-2015 годах, что является значимым на 1% уровне. Кроме того, однопроцентное увеличение пользователей дебетовых карт и темпов роста ВВП привело к 1,267% увеличению и 1,768% снижению соответственно розничных продаж в электронной коммерции, но имеет незначительную положительную связь с пользователями кредитных карт. Когда темпы роста ВВП являются зависимой переменной, то рост и пользователи дебетовых карт значительно положительно связаны, но когда розничные продажи являются зависимой переменной, то пользователи Интернета и розничные продажи значительно положительно связаны. Тест причинности Грейнджера подтвердил, что общий объем розничных продаж в электронной коммерции ( $x_1$ ), процент населения, использующего интернет ( $x_2$ ), количество кредитных ( $x_3$ ) и дебетовых карт ( $x_4$ ), темпы роста ВВП Индии ( $y$ ) в 2005-2020 гг. показали двунаправленную причинность, за исключением того, что  $X_2$  не является причиной Грейнджера для  $X_1$ , а  $x_2$  и  $x_4$  не имеют причинности.

Подобного рода исследование факторов, влияющих на развитие интернет-торговли было проведено в Турции и представлено в статье [15].

В анализе использовались ежемесячные данные за период с января 2010 года по декабрь 2014 года. Общее количество наблюдений составляет шестьдесят. Авторами был отобраны следующие переменные:

- ежемесячный объем операций электронной коммерции – внутренние онлайн-транзакции с использованием внутренних и международных карт в Турции (данные с сайта ВКМ);
- ежемесячный индекс потребительских цен по отношению к 2003 году (из базы данных Центрального банка Турции);
- ежемесячное общее количество пользователей Интернета в Турции (согласно данным с сайта IAB Turkey);
- ежемесячный объем импорта в млн. tl (из базы данных Центрального банка Турции);
- ежемесячный объем операций с кредитными картами - общий объем транзакций на внутреннем рынке (млн. TL), совершенных с использованием внутренних и международных карт (согласно данным Card Monitor 2014, опубликованным ВКМ).

В результате всех анализов было выявлено значительное влияние индекса потребительских цен, использования кредитных карт и импорта на электронную коммерцию. В то время как влияние импорта и использования кредитных карт на электронную коммерцию положительное, влияние индекса потребительских цен на электронную коммерцию отрицательное. Согласно анализу, в отличие от литературных данных, между электронной коммерцией и переменной интернет-пользователей не было обнаружено статистически значимой связи.

Также были проведены исследования, где интернет-торговли является не зависимой переменной, а фактором развития для других показателей.

Например, в работе [16] доказывается сильная взаимосвязь между ВВП и объемом электронной коммерцией по данным за период с 2004 по 2018 гг. по европейским странам: интернет-продажи напрямую влияют на экономический рост государства.

Также сильная взаимосвязь подтверждается в работе [17] между электронной коммерцией (объем транзакций в электронной коммерции), уровнем логистики (объем грузооборота) и ростом ВВП. На основании

данных Китая за 2014-2017 гг. авторами была построена модель линейной регрессии, коэффициент  $R^2$  которой составил 0,993, а значимость показателей была выше 0,05.

Исследования факторов, оказывающих влияние на развитие рынка интернет-торговли, производились и российскими авторами. Рассмотрим также несколько представленных работ.

В статье [18] автором на основании данных по субъектам РФ за 2011-2017 гг. модель регрессии для переменных:

- доля рынка электронной коммерции в обороте розничной торговли (зависимая переменная);
- доля домашних хозяйств, имеющих доступ к сети интернет, в общем числе домашних хозяйств;
- доля домашних хозяйств, использующих для выхода в сеть Интернет мобильные телефоны или смартфоны;
- реальные располагаемые денежные доходы по Российской Федерации.

В исследовании была подтверждена положительная взаимосвязь доли рынка электронной коммерции и долей домашних хозяйств с доступом в Интернет, и отрицательная взаимосвязь между объемом интернет-торговли и реальными располагаемыми доходами.

Рассмотрим также похожую работу [19], объектом которой выступают данные по федеральным округам РФ за 2014 год, согласно Федеральной службы государственной статистики. В качестве зависимой переменной выступает объем розничной торговли в сети интернет (млрд. руб.). Детерминантами были выбраны следующие показатели:

- проникновение интернета (%);
- индекс доступности интернета;
- коэффициент официальной ликвидации организаций;
- доля автомобильных дорог общего пользования, не отвечающих нормативным требованиям (%);

- численность населения с доходами ниже прожиточного минимума (%).

Автором статьи были построены несколько регрессионных моделей, в ходе анализа которых постепенно исключались наименее значимые факторы. Таким образом, в результате исследования были сделаны следующие выводы:

- доступность интернета в регионе (как фактор, учитывающий как доходы населения, так и среднюю стоимость доступа к сети) оказывает наиболее заметное влияние на развитие розничной торговли в интернете;

- прямая связь оборота торговли в сети и ликвидации организаций свидетельствует об увеличении продаж по причине ухода предпринимателей в интернет с целью сокрытия доходов и ухода от налогового бремени;

- состояние дорог в регионе не оказывает значительного влияния на объём продаж, как предполагалось исследователями;

- проникновение интернета в регион оказывает влияние на объём продаж с определённым временным лагом.

Дополнительно в работе [20] рассматривается в большой степени социально-демографические факторы, которые могут оказывать влияние на рынок электронной коммерции. Для анализа были отобраны следующие переменные:

Y - объём рынка электронной торговли, млрд.руб.;

X1 - оборот розничной торговли, трлн.руб.;

X2 - индексы потребительский цен на товары и услуги, %;

X3 - среднемесячная номинальная заработная плата, руб.;

X4 - среднедушевые денежные доходы населения, руб.;

X5 – индекс потребительской уверенности, %.

В ходе корреляционного анализа авторами были отобраны только два показателя: оборот розничной торговли и среднедушевые денежные доходы.

Регрессионный анализ показал, что 99% вариации уровня объемов рынка электронной торговли Российской Федерации зависимы от факторов, включенных в исследуемую модель.

Проанализировав уравнение регрессии, автором делаются следующие выводы: с изменением оборота розничной торговли объем рынка электронной торговли уменьшится на 2,7 млрд. руб., а с ростом среднедушевых денежных доходов населения объем электронной торговли увеличится на 0,01 млрд. руб. Зависимость от анализируемых факторов очевидна: низкие цены и достаточные доходы влияют на возможность людей покупать товары «у дома», что, в свою очередь, приводит к нежеланию обращаться к интернет-магазинам (хоть бы в силу ожидания доставки товаров).

Также российским автором Коротких В.В. в статье [21] было проведено исследование факторов на российском рынке интернет-торговли по открытым данным за 2011-2021 гг. Для анализа исследователь выделил следующие показатели, представленные в таблице 1.

Таблица 1 – Выделенные показатели для анализа из работы Коротких В.В. [21]

Показатель	Источник
Доля онлайн-продаж в розничных продажах	Федеральная служба государственной статистики РФ; Ассоциация компаний интернет-торговли (АКИТ); ПАО «Сбербанк»
Темп прироста розничных онлайн-продаж в текущем году	Рассчитаны автором на основе данных Федеральной службы Государственной статистики РФ
Темп прироста розничных онлайн-продаж в предыдущем году	
Проникновение онлайн-торговли в России, оцененное как доля городских жителей в возрасте 16-55 лет, покупающие онлайн чаще раза в год	Федеральная служба государственной статистики РФ; Международный союз электросвязи (International Telecommunication Union)
Индекс эффективности логистики	Всемирный банк (WorldBank); Всемирный почтовый союз (Universal Postal Union)
Индекс человеческого капитала (учитывает уровень грамотности взрослого населения; совокупный валовой коэффициент охвата начальным, средним и высшим образованием; ожидаемую и среднюю продолжительность обучения)	Отдел государственного управления и управления развитием департамента по экономическим и социальным вопросам ООН

### Окончание таблицы 1

Индекс качества онлайн-обслуживания (вычисляется на основе анкетного опроса, включающего 148 закрытых вопросов)	
Доля онлайн-заказов, получаемых в пункте выдачи	Яндекс.Маркет; ООО «ГфК-Русь» (Growth from Knowledge Rus)
Переменная для учета фиксированных эффектов в динамике электронной коммерции, связанных с COVID-19	переменная, принимающая значение равное «1» в 2020 г. и «0» – для остальных лет

Автором было построено четыре регрессионные модели с разным составом факторов по переменной Sales.

Коэффициент R<sup>2</sup> по построенным моделям составил от 73,4% до 99,6%. Исследователем был сделан следующий вывод: статистически значимое положительное влияния на долю онлайн-торговли в структуре розничной оказали проникновение онлайн-торговли в РФ, уровень развития логистических систем, индекс человеческого капитала и пандемия COVID-19, что свидетельствует об ускоренном росте спроса и предложения, в частности товаров повседневного потребления (как правило, продукция легкой и пищевой промышленности), в период изоляции и после него. Доля онлайн-заказов, получаемых в пункте выдачи, положительно сказывается на динамике исследуемой переменной только при исключении из спецификации переменных, характеризующих проникновение онлайн-торговли в РФ и уровень развития логистических систем. Это свидетельствует о том, что открытие дополнительных пунктов выдачи онлайн-заказов без сопутствующего развития логистических систем не будет способствовать росту объемов электронной коммерции в РФ. В настоящее время на 1 адрес уже приходится более 2 пунктов выдачи онлайн-заказов.

На основании рассмотренных исследований, касающихся изучения факторов, воздействующих на изменения интернет-торговли, можно структурировать предполагаемые показатели. Переменные можно условно разделить на четыре группы: зависимые переменные, экономические факторы, социально-демографические факторы.

Общий список показателей, рассматриваемые различными авторами, упомянутыми в данной работе, представлен в таблице В.1.

Таким образом, рассмотрим и подберем также по группам показатели, которые могут оказывать влияние на развитие интернет-торговли в России.

## **2. Данные и методы**

### **2.1. Описание исходных данных**

Для проведения данного исследования, а именно выявления возможных факторов, оказывающих влияние на интернет-торговлю в России, были отобраны различные показатели по результатам литературного обзора за период с 2016 по 2021 год по субъектам РФ (было отобрано 81 субъект РФ) из различных источников (преимущественно ФСС (Росстат)). В общей сложности по каждому показателю было проведено 486 наблюдений (за 6 лет по 81 субъекту). Все переменные были разделены на пять групп – зависимая переменная ( $y$ ), социально-демографические факторы (соц  $x$ ), экономические факторы (эк  $x$ ), технологические факторы (тех  $x$ ), логистические факторы (лог  $x$ ). Список всех показателей представлен в таблице В.2.

В качестве зависимой переменной предлагается рассмотреть показатель продаж через Интернет в виде доли от розничных продаж по регионам.

На основании этого можно предложить следующую гипотезу исследования: в совокупности набор различных социально-демографических, (кроме показателя  $progror$ ), экономических, технологических и логистических факторов положительно связана с изменением доли продаж через Интернет в России.

Таким образом, для исследования был отобран 20 показателей по различным группам факторов за 2016-2021 гг. по 81 субъекту РФ.

### **2.2. Методы и модели исследования**

Для проверки гипотез будет использоваться множественный регрессионный анализ, основанный на методе наименьших квадратов (формула 1).

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + u_i, \quad (1)$$

где:

$\beta_0$  – константа;

$y_i$  – зависимая переменная;

$\beta_k$  – параметр при переменной  $x_k$ ;

$u_i$  – случайная ошибка.

Исследуемая выборка содержит данные о нескольких объектах, каждый из которых наблюдается в течение 2016-2021 гг, поэтому их можно охарактеризовать как панельные данные. Таким образом, также целесообразно построить и рассмотреть модели с фиксированными и/или случайными эффектами.

Для сглаживания показателей и достижения линейности по параметрам помимо стандартной линейной модели параллельно будут рассмотрены модели различных форм: линейно-логарифмические ( $y-\ln x$ ), логарифмические ( $\ln y-\ln x$ ), логарифмически-линейные ( $\ln y-x$ ) модели.

Для отбора факторов, а также проверки на мультиколлинеарность будет применен корреляционный анализ и рассчитан показатель VIF (Variance Inflation Factor) для каждой анализируемой модели, чтобы избежать зависимостей между переменными.

Для выбора наилучшей модели будут использованы следующие информационные критерии: R-squared, F-statistic, P-value, AIC, BIC.

Помимо этого, необходимо провести тестирование, проверки нарушений предпосылок моделей по различным методам.

1. Проверка линейности: RESET тест (тест Вальда), частичные остаточные графики.
2. Гетероскедастичность: Тест Бреуша-Пагана, тест Голдфелда-Кувндта, графики рассеивания остатков.
3. Автокорреляция остатков: Статистика Дарбина-Уотсона.
4. Нормальность распределения остатков: Q-Q график (Quantile-Quantile plot).



По рассчитанным параметрам будет отобрана модель, которая наилучшим образом описывает изменения продаж в Интернете в Российской Федерации.

### 3. Результаты анализа

#### 3.1. Результаты анализа описательной статистики

На первом этапе исследования необходимо провести предварительный анализ данных. В таблице В.3 представлены некоторые основные результаты описательной статистики для всех 81 субъектов РФ за 2016-2021 гг. по выделенным показателям за исключением covid, так как это фиктивная переменная.

Отобразим данные для всех 81 субъектов РФ за 2016-2021 гг. по зависимой переменной eshare по разделению фиктивной переменной covid, а также возведем в логарифм переменную eshare и также отобразим на графиках.

Отобразим на рисунке 1 графики распределения частот в процентах по переменным eshare и логарифмированной ln\_eshare с разбиением по фиктивной переменной covid.

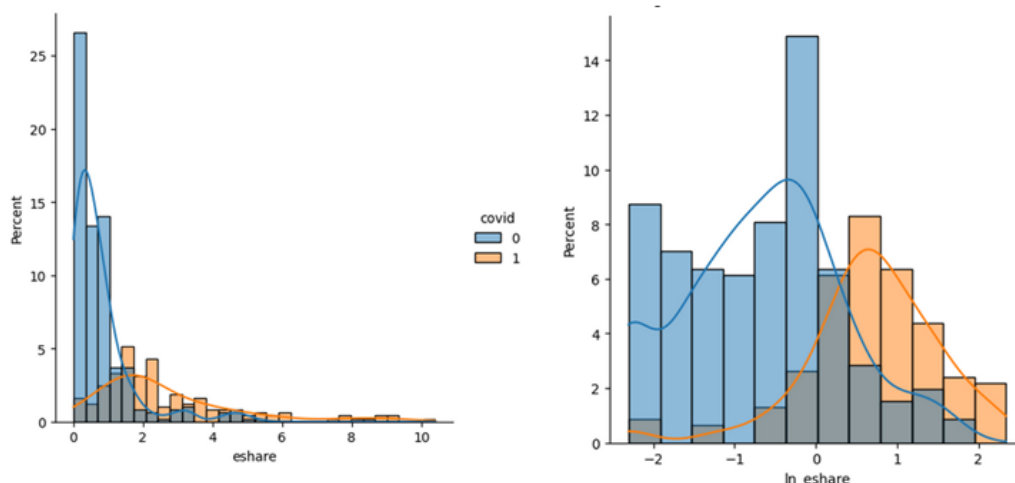


Рисунок 1. – График распределения по переменным eshare и ln\_eshare

На графиках распределения можно наблюдать, что среднее и максимальное значения в период коронавируса больше, чем в период до него. Также стандартное отклонение выше во время эпидемии. Помимо этого,

стоит отметить, что логарифмирование переменной *eshare* приводит к появлению отрицательных значений.

Также построим графики ящиков с усами по рассматриваемым переменным, где отобразим основные статистические значения распределения данных.

На рисунке 2 отобразим график по переменным *eshare* и *ln\_eshare*.

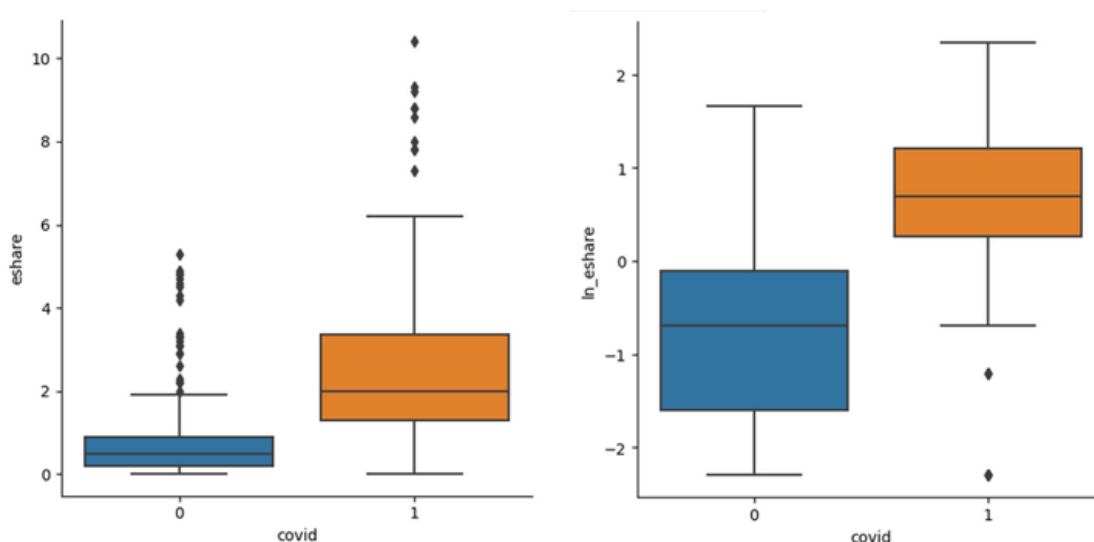


Рисунок 2. – Графики «ящик с усами» по переменным *eshare* и *ln\_eshare*

На графике *eshare* видно, что имеется некоторое количество выбросов, которые во многом превышают средние значения по группам. Также можно наблюдать, что медианное значение в период коронавируса выше, чем до него. Помимо этого, разброс значений (межквартильный размах) в период коронавируса гораздо больше, чем до вспышки. Максимальное значение также выше в случае коронавируса. Можно сделать предположение, что вспышка коронавируса все же повлияла на увеличение доли продаж в Интернете, так как значения по выборке гораздо выше, чем в выборке без эпидемии. В случае с переменной *ln\_eshare* тенденция повторяется за исключением межквартильного размаха: в период до коронавируса межквартильный размах больше.

Рассмотрим также описательную статистику для 2016 и 2021 года в динамике. Данные представлены в таблицах В.4 и В.5. Также отобразим динамику распределения по 2016 и 2021 на рисунке 3.

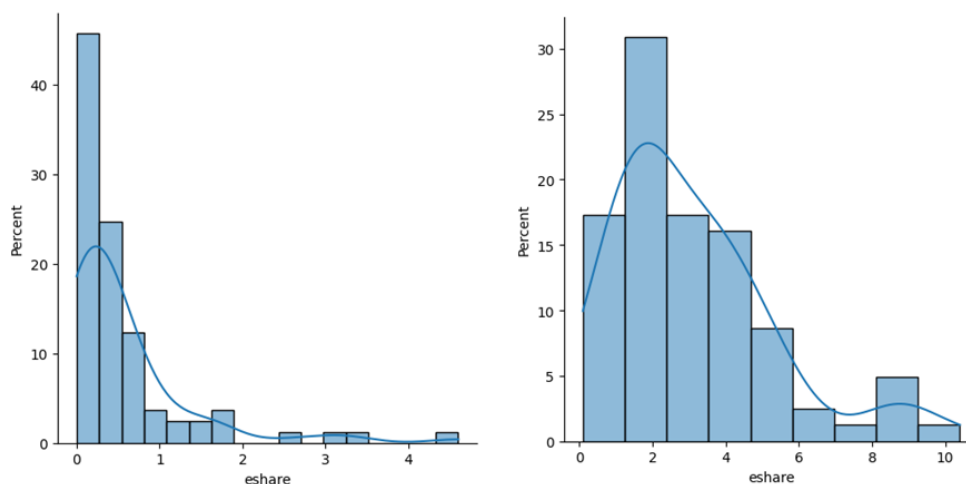


Рисунок 3. – Графики распределения по показателю eshare для 2016 и 2021 гг. соответственно

Со временем прослеживается рост среднего значения распределения, а также повышение максимального значения. Так в 2016 году среднее значение по распределению показателя eshare составляло 0,56%, а максимум – 4,6%, а в 2021 году среднее значение составило 3,11%, а максимум – 10,4%. Вместе с этим возросло и стандартное отклонение с 0,8% до 2,24%.

Проведем корреляционный анализ и отобразим корреляционные матрицы по переменным eshare и ln\_eshare по выделенным группам факторов.

Для начала рассмотрим корреляцию между переменными eshare и ln\_eshare и социально-демографической группой факторов. На **рисунке 4** отображена их корреляционная матрица.

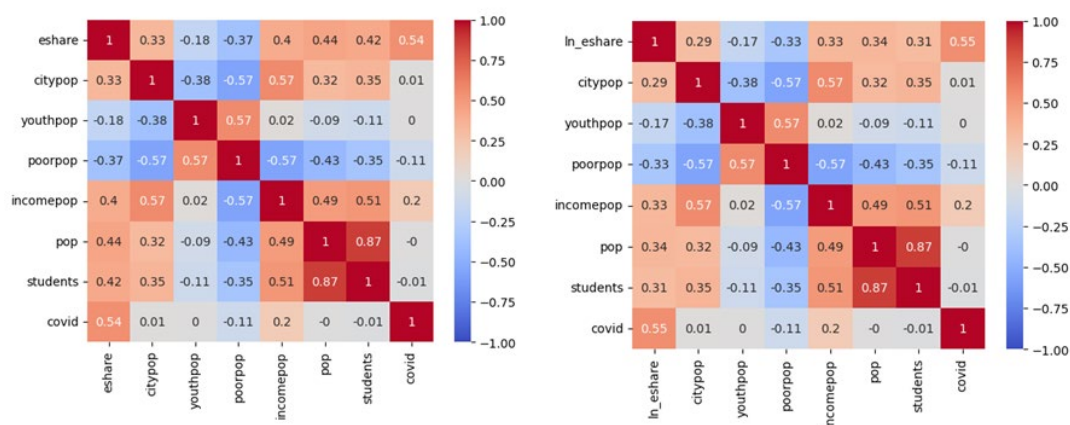


Рисунок 4 – Корреляционная матрица для социально-демографических факторов

На матрице не наблюдается значимых корреляционных связей за исключением сильной зависимости между показателями students и pop на уровне 87%.

Далее рассмотрим корреляционные матрицы с группой экономических факторов по переменным eshare, ln\_eshare (рисунок 5).

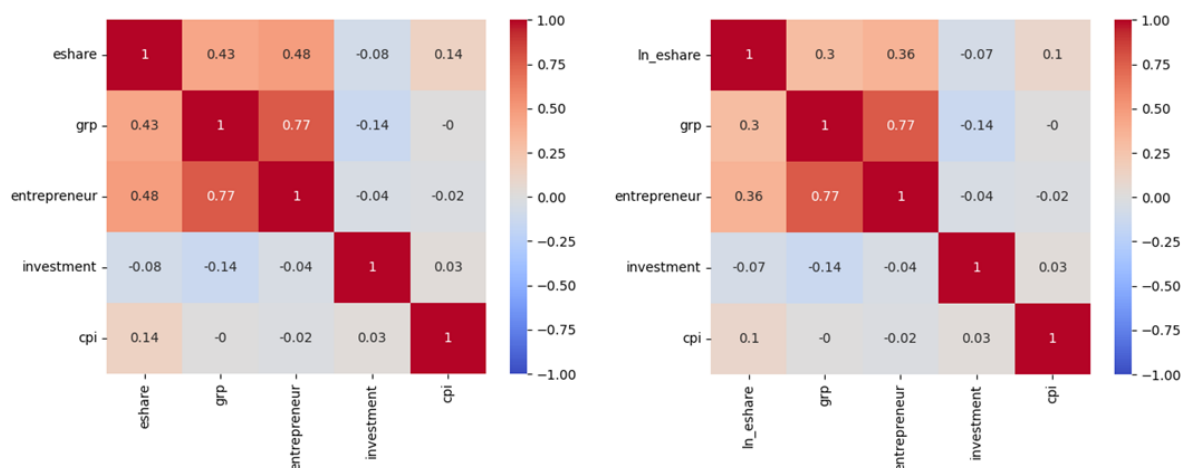


Рисунок 5 – Корреляционная матрица для экономических факторов

Также на всех матрицах наблюдается средняя связь между grp и entrepreneur на уровне 77%.

Далее рассмотрим корреляционные матрицы с группой технологических факторов по переменным eshare, ln\_eshare (рисунок 6).

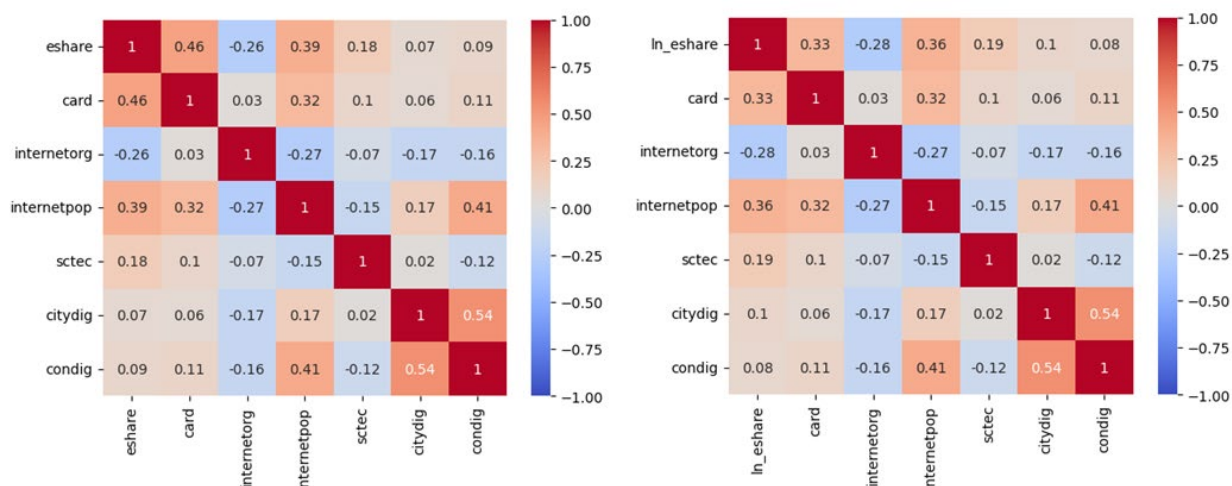


Рисунок 6 – Корреляционная матрица для технологических факторов

Между зависимыми переменными и технологическими факторами не наблюдается сильных зависимостей.

В заключение также отобразим корреляционные матрицы по группе логистических показателей по переменным *eshare*, *ln\_eshare* (рисунок 7).

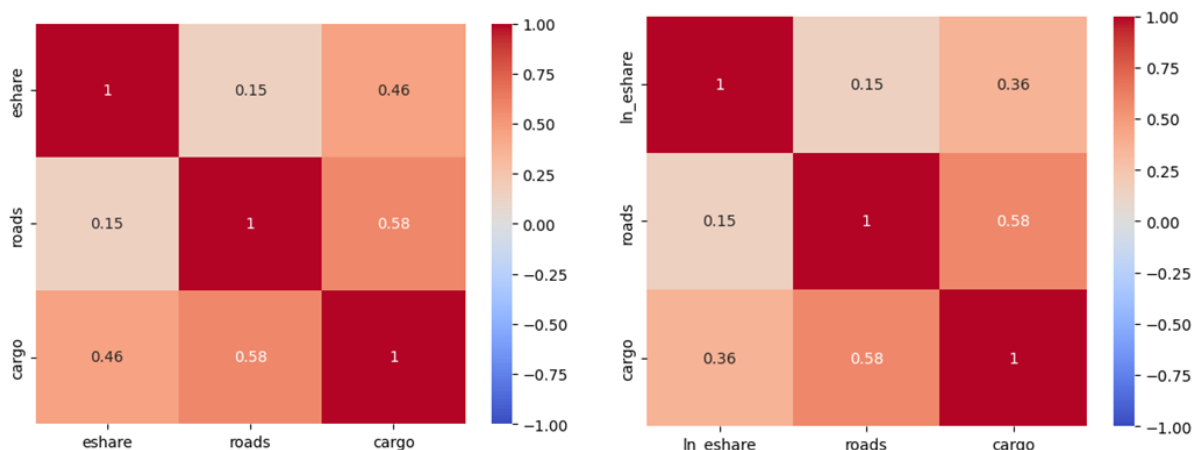


Рисунок 7 – Корреляционная матрица для логистических факторов

Совокупные матрицы корреляции по всем группам факторов показали следующие сильные связи:

- сильная положительная корреляция между *students* и *pop* (87%);
- сильная положительная корреляция между *entrepreneur* и *pop* (95%);
- сильная положительная корреляция между *grp* и *pop* (86%);
- сильная положительная корреляция между *grp* и *students* (80%);
- сильная положительная корреляция между *entrepreneur* и *students* (80%);
- сильная положительная корреляция между *card* и *pop* (90%);
- сильная положительная корреляция между *card* и *students* (90%);
- сильная положительная корреляция между *card* и *grp* (96%).

На данном этапе логичнее всего исключить из рассмотрения переменную *pop*, так как она связана с показателями, которые содержат значения численности человек. Также можно сделать выводы по тому, что в общей регрессионной модели будет наблюдаться мультиколлинеарность по показателям *students*, *grp*, *entrepreneur*, *card*.

Далее по рассматриваемым гипотезам построим модели и проведем их оценку согласно различным тестам. В качестве зависимой переменной будет выступать показатель *eshare*.

### **3.2. Анализ результатов эконометрического моделирования**

Для проверки наших гипотез было построено шесть моделей обычной множественной регрессии и проведены тесты для определения всех предпосылок применения регрессии. Результаты построения всех моделей представлены в таблице В.5. Также в таблице В.6. представлены расчеты параметров VIF.

Таким образом, для улучшения качества модели были выполнены следующие шаги.

1. Изменена форма представления, были возведены в логарифм как зависимая переменная, так и регрессоры.

2. Из данных удалены нулевые значения по показателю eshare, а также выбросы по следующим субъектам за 2016-2021гг.: Республика Ингушетия, Камчатский край, Магаданская область, г. Севастополь, Забайкальский край, Республика Бурятия, Республика Дагестан, Республика Калмыкия, Республика Северная Осетия, Республика Карачаево-Черкесия, Чеченская Республика, Республика Саха (Якутия), Еврейская Автономная область, Новосибирская область, Томская область, Калининградская область, Чувашская Республика, Республика Крым, г. Москва, г. Санкт-Петербург, Тюменская область, Республика Алтай, Республика Башкортостан, Республика Тыва, Московская область, Республика Адыгея. Таким образом, данные сократились до 330 наблюдений.

3. Избавились от мультиколлинеарности путем удаления коррелированных между собой переменных и незначимых показателей. Итоговый список анализируемых показателей стал выглядеть следующим образом: `ln_citypop`, `ln_youthpop`, `ln_incomepop`, `ln_students`, `covid`, `ln_investment`, `ln_cpi`, `ln_internetorg`, `ln_internetpop`, `ln_sctec`, `ln_citydig`, `ln_roads`.

Проведем тестирование на предпосылки регрессии.

После всех преобразований графики расстояния Кука и нормализованных остатков стали выглядеть следующим образом (рисунок 8).

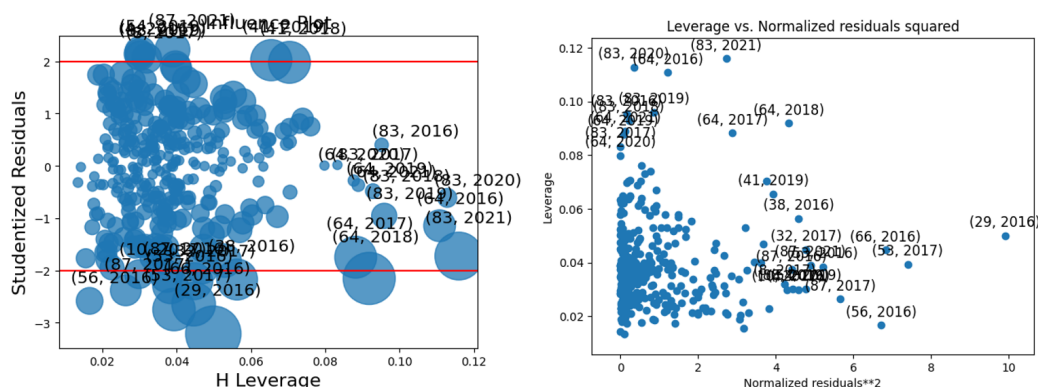


Рисунок 8 – Графики расстояния Кука и нормализованных остатков для линейной модели №6

Проведем тестирование на наличие гетероскедастичности. На рисунке 9 представлен график остатков модели по зависимой переменной  $\ln\_eshare$ .

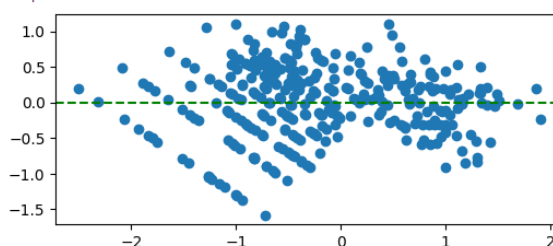


Рисунок 9 – График распределения остатков по переменной  $\ln\_eshare$  для модели №6

Согласно тесту Бреуша-Пагана, присутствует гетероскедастичность остатков, что может повлиять на эффективность коэффициентов переменных (рисунок 10).

```
[('X square statistic', 40.029213781583806),
 ('p-value', 7.111036801376752e-05),
 ('f-value', 3.6467066603059086),
 ('f p-value', 3.8442312662689754e-05)]
```

Рисунок 10 – Результаты теста Бреуша-Пагана для модели №6

Однако, согласно тесту Голдфелда-Кванта, остатки модели гомоскедастичны (рисунок 11). Поэтому нельзя однозначно утверждать, что в модели №6 присутствует гетероскедастичность.

```
[('Goldfeld Quandt statistic', 1.2824972334914222),  
 ('p-value', 0.06309786388380094)]
```

Рисунок 11 – Результаты теста Голдфелда-Квандта для модели №6

Отобразим на рисунках В.11 и В.12 графики частичной регрессии и частичных остатков, а также проведем тесты на спецификацию (линейность) модели.

На графиках не наблюдается выбросов, сдвигающих линию регрессии в неверном направлении.

Согласно Linear Rainbow тесту, спецификацию модели можно считать приемлемой, между переменными присутствует линейная зависимость (рисунок 12).

```
[('Rainbow test statistic', 1.0509371481290037),  
 ('p - value', 0.37839880784492186)]
```

Рисунок 12 – Результаты Linear Rainbow Test для модели №6

Согласно RESET тесту, спецификацию модели нельзя однозначно считать приемлемой, но значение близко к тому, чтобы принять нулевую гипотезу теста (рисунок 13).

```
<Wald test (chi2): statistic=10.707052095881355, p-value=0.030061316315604377, df_denom=4>
```

Рисунок 13 – Результаты RESET-теста для модели №6

Согласно графикам на рисунке 14, можно подтвердить нормальность распределения остатков модели.

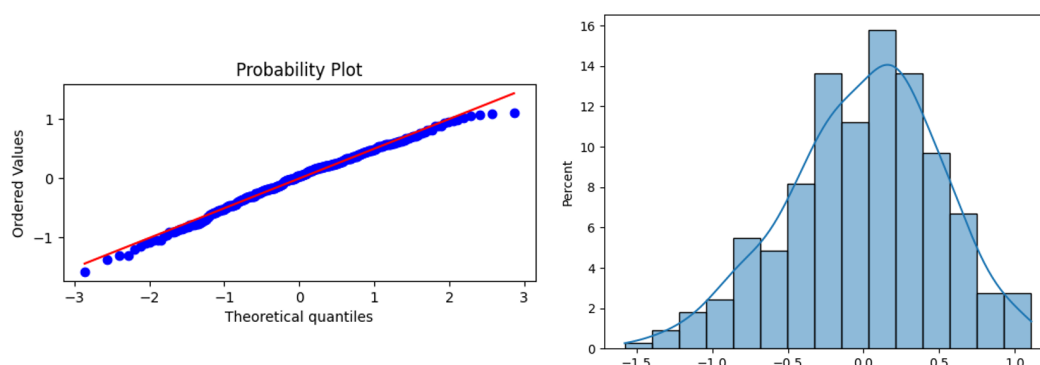


Рисунок 14 – Графики нормального распределения остатков для модели №6

Также необходимо отметить, что в модели присутствует небольшая автокоррелированность остатков. Значение теста Дарбина-Уотсона



находиться почти у порогового значения зоны неопределенности от 1 до 2 (рисунок 15).

0.9448918912647051

#### Рисунок 15 – Результаты теста Дарбина-Уотсона для модели №6

В результате анализа и тестирования моделей на выполнение предпосылок итоговая модель №6 имеет следующие упущения:

- 1) остатки модели гетероскедастичны и автокоррелированы, поэтому коэффициенты переменных нельзя считать достаточно состоятельными;
- 2) модель №6 соответствует предпосылкам о нормальности распределения остатков, описана линейной зависимостью (согласно Linear Rainbow Test и графика частичной регрессии и остатков) и не имеет мультиколлинеарности.

Таким образом, модель №6 соответствует трем из пяти предпосылок, и оценки коэффициентов при переменных-предикторов нельзя считать полностью состоятельными. Для дополнительного анализа были построены модели со случайными и фиксированными эффектами (рисунок В.13). Но так как многие показатели потеряли свою значимость было принято решение остановиться на обычной МНК модели.

#### **4. Обсуждение результатов**

Проанализируем полученные коэффициенты при независимых переменных из модели №6.

Наиболее значимыми переменными можно считать следующие показатели:  $\ln\_citypop$ ,  $\ln\_youthpop$ ,  $\ln\_incomepop$ ,  $covid$ ,  $\ln\_internetorg$ ,  $\ln\_roads$ . Вторыми по значимости можно считать следующие показатели:  $\ln\_citydig$ ,  $\ln\_cpi$ ,  $\ln\_sctec$ ,  $\ln\_internetpop$ ,  $\ln\_investment$ ,  $\ln\_students$ .

При этом  $youthpop$  и  $\ln\_cpi$  имеют отрицательную зависимость с  $\ln\_eshare$ , тогда как остальные переменные – положительную.

Таким образом, интерпретируем полученный результат и сравним с исследованиями других авторов.

Показатель доли городского населения ( $\ln\_citypop$ ) положительно влияет на долю интернет-продаж. Данный результат получили исследователи в работе [9]. В городах доступность Интернета гораздо выше, чем в сельской местности, а также больше количество сервисов по доставке и пунктов выдачи. Таким образом, городскому населению легче совершать покупки через интернет. Данный результат оказался ожидаемым.

Положительная зависимость уровня образования ( $\ln\_students$ ) и темпов роста продаж через Интернет подтверждает в работах A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat [9] и Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega [13]. Чем выше образование человека, тем скорее всего выше его навыки пользования Интернетом и выше вероятность совершения его покупок через Интернет. Данный результат оказался ожидаемым.

Наибольшее упоминание в литературных источниках получил показатель, характеризующий использование интернета населением, количество его активных пользователей. Положительный эффект был отмечен в работе [9], [10], [14], [15], [11], [18], [19]. Похожий результат был получен и в данном исследовании: чем больше пользователей Интернет, тем больший объем интернет-покупок совершается. Данный результат оказался ожидаемым.

Значимость логистики была отмечена в статьях [9] и [21]. В нашем случае, положительное влияние характеризуется через взаимосвязь интернет-продаж и протяженностью автомобильных дорог с покрытием. Чем больше протяженность хорошо асфальтированных дорог, тем больше регионов охватывается и больше возможностей для организации доставок товаров через Интернет

Положительная взаимосвязь между доходами населения ( $\ln\_incomeror$ ) и продажами через интернет подтверждается в работах [11] и [20]. Чем выше доход у человека, тем больше покупок он может совершить.

Отрицательная зависимость между индексом потребительских цен и интернет-торговлей отмечается также в работе Filiz Kadı, Canan Peker [15]. Чем выше цены на товары, тем менее охотнее совершаются покупки через Интернет.

Также в работе [11] отмечается положительное влияние процента от ВВП, посвященный исследованиям и расходы на развитие ( $\ln\_sc tec$  и  $\ln\_investment$ ) на темпы роста интернет-торговли и цифровизации в целом. Эффект цифровизации ( $\ln\_citydig$ ) был отмечен во всех рассмотренных ранее исследованиях. В свою очередь это оказывает влияние на долю организаций, использующих в своей деятельности Интернет. Чем выше процент таких компаний, тем скорее всего будет больше предложений в Интернете от них (если это торговое предприятие).

Ожидаемым результатом оказалось воздействие пандемии коронавируса в 2020-2021 гг. Но данное событие нельзя считать фактором развития интернет-торговли в России, скорее всего это было лишь толчком для ускорения этого роста. Такое же заключение отмечается в работах [21] и [14].

Совершенно неожиданным результатом стала обратнопропорциональная зависимость между показателем доли молодого населения в трудоспособном возрасте и продаж через Интернет. Скорее всего это связано с пандемией коронавируса, когда пожилому населению нельзя было выходить на улицу в целях сохранения своего здоровья, поэтому большинство покупок приходилось совершать через Интернет, в том числе и продукты. Поэтому данный результат нельзя считать фактором, драйвером развития интернет-торговли.

Таким образом, выдвинутая гипотеза исследования считается подтвержденной частично, так как не все отобранные показатели оказались статистически значимыми. Но было доказано, что переменные разного характера (социально-демографические, экономические, технологические, логистические) оказывают влияние на динамику развития интернет-торговли.

## **ЗАКЛЮЧЕНИЕ**

В данной работе была рассмотрена динамика рынка интернет-торговли как в России, так и в странах мира. На основании различных исследований делается вывод о положительном развитии розничных интернет-продаж в России, изучаются возможные показатели и факторы, которые могли оказать существенных рост.

Целью данного исследования было изучить влияние различных социально-демографических, технологических, экономических и логистических факторов на рост электронной коммерции на основании изучения открытых данных по субъектам РФ. Было обнаружено, что доля розничных продаж через интернет положительно связанасо следующими факторами: доля молодого населения, доли компаний и населения, использующих интернет, пандемия коронавируса, среднедушевые денежные доходы населения, индекс потребительских цен, доля городского населения, протяженность автомобильных дорог с покрытием, цифровизация телефонной сети в городах, уровень образования населения, доли инвестиций в исследование и развитие технологий.

Все результаты, кроме взаимосвязи с долей молодого населения, были ожидаемыми. Исследование показало, что существует отрицательное влияние процента молодого населения на продажи через Интернет, но данный показатель нельзя считать полностью достоверным, так как в некоторых построенных моделях он оказывался статистически незначимым.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Ермилова С.В. Влияние пандемии на развитие электронной торговли в России / С.В. Ермилова // Российский экономический вестник. 2022. Том 5. № 2. С. 199-204.
2. Шайдуллина В.К. Электронная торговля и перспективы ее развития в мировой экономике / В.К. Шайдуллина // Вестник университета. Экономика: проблемы, решения и перспективы. 2019. № 3. С. 114-119.
3. Кузьмичёва И.А. Перспективы развития рынка интернет-торговли / И.А. Кузьмичёва, Л.А. Сахарова, А.О. Пугина // Вектор науки Тольяттинского Государственного Университета. Серия: экономика и управление. 2021. № 2(45). С. 27-33.
4. Слободняк И.А. Оценка статистических данных о состоянии рынка электронной торговли с использованием приемов экономического анализа / И.А. Слободняк, П.В. Антипина // Инвестиционный, финансовый и управленческий анализ. 2020. Том 2. № 1. С. 150-160.
5. Седых И. А. Рынок интернет-торговли в РФ //М: Национальный исследовательский университет Высшая школа экономики, Центр развития. URL: [https://www.rvc.ru/upload/iblock/0ee/e-commerce\\_resume\\_vk\\_part1\\_v1.pdf](https://www.rvc.ru/upload/iblock/0ee/e-commerce_resume_vk_part1_v1.pdf) (дата обращения: 20.10. 2022).
6. Сидоров Н.В. Особенности рынка электронной коммерции в России / Н.В. Сидоров, Э.Ш. Шаймиева // СИНЕРГИЯ НАУК. 2017. № 12. С. 407-414.
7. Тюрин Д.В. Специфика развития онлайн-торговли под влиянием потребительского поведения интернет-пользователей / Д.В. Тюрин, А.О. Волкова // Вестник РГГУ. Серия «Экономика. Управление. Право». 2021. No 3 (Ч. 2). С. 158–171.
8. Давий А.О. Барьеры и драйверы при совершении интернет-покупок в России: результаты эмпирического исследования / А.О. Давий, В.А. Ребязина, М.М. Смирнова // Вестник СПбГУ. Менеджмент. 2018. Том 17. № 1.

9. Waseem A. FACTORS AFFECTING E-COMMERCE POTENTIAL OF ANY COUNTRY USING MULTIPLE REGRESSION ANALYSIS / A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat // Journal of internet banking and commerce, August 2019, vol. 24, no. 2.

10. Dumičić K. Regression analysis of e-commerce: focusing the selected EU candidates and the EU countries / K. Dumičić, B. Žmuk, B. Novkovska // Conference: Proceedings of the 7th International Scientific Conference "Knowledge and Business Challenge of Globalisation in 2017", FKPV, Celje (Editors: Kovač, T.; Cingula, M.)At: Celje, Slovenia, 2017, P. 156-164.

11. Lucero Ortiz A. E-commerce Development in Europe: A Panel Data Analysis 2003–2017 / Lucero Ortiz A., Rodríguez J. C., Gómez M // Business Administration and Management. 2020. XXIII, 4. P. 89-101."

12. Rodríguez-Ardura I. Factors Influencing the Evolution of Electronic Commerce: An Empirical Analysis in a Developed Market Economy / I. Rodríguez-Ardura, A. Meseguer-Artola, J. Vilaseca-Requena // Journal of Theoretical and Applied Electronic Commerce Research. 2008. VOL 3. ISSUE 2. P 18-29.

13. Fernández-Bonilla F. E-commerce in Spain: Determining factors and the importance of the e-trust / Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega // Telecommunications Policy. 2022. Volume 46. Issue 1. P. 1-12.

14. Dr. Debesh Bhowmik. India's Retail Sales of ECommerce: An Econometric Analysis / Dr. Debesh Bhowmik // International Journal of Scientific and Research Publications. 2017. Volume 7. Issue 2. P. 13-21.

15. Filiz Kadı. Analyzing the Factors Affecting E-Commerce in Turkey / Filiz Kadı, Canan Peker // International Journal of Management, Accounting and Economics. 2015. Volume 2. Issue 11. P. 1319-1339.

16. Parishev A. E-commerce impact on economic growth / A. Parishev, G. Hristovski, P. Jolakovski, V. Stojkoski // Conference: 1st international scientific conference "Economic and Business Trends Shaping the Future" At: Skopje, North Macedonia, 2020, P. 188-198.

17. Tang X. Design and analysis of e-commerce and modern logistics for regional economic integration in wireless networks / X. Tang, G. Wang // EURASIP Journal on Wireless Communications and Networking, 2020, № 208.

18. Мотина В. Г. Факторный анализ динамики рынка интернет-торговли / В. Г. Мотина // Актуальные проблемы перспективы развития экономики : Труды XVII Международной научно-практической конференции, Симферополь - Гурзуф, 18–20 октября 2018 года / Под редакцией Н.В. Апатовой. – Симферополь - Гурзуф: ИП Зуева Т.В., 2018. – С. 270-273.

19. Видякин А. Н. Факторы развития интернет-торговли в России / А. Н. Видякин // Вестник Поволжского государственного технологического университета. Серия: Экономика и управление. – 2017. – № 1(33). – С. 51-60.

20. Михайленко А.А. АНАЛИЗ РОССИЙСКОГО РЫНКА ЭЛЕКТРОННОЙ ТОРГОВЛИ / А.А. Михайленко // ЭЛЕКТРОННЫЙ НАУЧНЫЙ ЖУРНАЛ «ВЕКТОР ЭКОНОМИКИ». 2020. №1.

21 Коротких В.В. СТАТИСТИЧЕСКИЙ АНАЛИЗ ФАКТОРОВ РАЗВИТИЯ ЭЛЕКТРОННОЙ КОММЕРЦИИ ПО ОТКРЫТЫМ ДАННЫМ / В.В. Коротких // СОВРЕМЕННАЯ ЭКОНОМИКА: ПРОБЛЕМЫ И РЕШЕНИЯ. 2022. № 8 (152). С. 71-87.

**ПРИЛОЖЕНИЕ А.**  
**Исходные данные исследования**



## ПРИЛОЖЕНИЕ Б.

### Код исследования

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import statsmodels.api as sm
import sklearn as sk
import scipy as sp
from statsmodels.compat import lzip
from statsmodels.stats.outliers_influence import variance_inflation_factor
from linearmodels.panel import PooledOLS
from linearmodels.panel import BetweenOLS
from linearmodels.panel import RandomEffects
from linearmodels.panel import PanelOLS
from linearmodels.panel import FirstDifferenceOLS
from linearmodels.panel import compare
```

```
pip install linearmodels
```

Рисунок Б.1 – Импорт необходимых библиотек

```
ecom = pd.read_excel('ecom16-21.xlsx')
elog = pd.read_excel('elog2.xlsx')
```

Рисунок Б.2 – Импорт датасета с подготовленными данными

### 3. Преобразование столбцов с годами и кодами субъектов РФ в индексы

```
year = pd.Categorical(ecom.year)
ecom = ecom.set_index(["cod", "year"])
ecom
```

Рисунок Б.3 – Преобразование столбцов с годами и кодами субъектов РФ в

### индексы

```
] f=pd.DataFrame(round(ecom.describe(percentiles=[0.1, 0.3,0.5, 0.7,0.9]),2))
#f.to_excel("Описательная статистика.xlsx")
f
```

	eshare	citypop	youthpop	poorpop	incomepop	pop	students	covid	grp	entrepreneur	investment	cpi	card	internetorg	internetpop	sctec	citydig	condig	roads	cargo
count	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00	486.00
mean	1.38	70.39	75.22	14.52	29654.60	1808924.70	51081.44	0.33	1140210.84	44736.52	18.85	104.58	3469287.12	78.72	81.11	19.32	93.86	85.66	13340.66	2253.19
std	1.66	13.09	4.26	5.04	10182.19	1820072.86	86937.63	0.47	2413590.31	51276.79	7.79	2.02	6413237.86	11.40	5.86	5.53	8.49	16.51	8752.60	2561.51
min	0.00	29.00	68.70	5.00	14963.00	138401.00	1947.00	0.00	47434.90	2434.00	3.50	100.84	184566.00	29.00	65.30	6.70	44.70	22.40	855.86	9.90
10%	0.10	53.10	70.90	8.85	20631.50	465844.00	7866.50	0.00	153425.30	11587.00	8.90	102.39	539955.50	61.20	73.75	11.95	82.55	61.50	3539.36	185.46
30%	0.40	66.25	72.50	12.20	24156.50	865581.50	18629.50	0.00	315598.50	18578.50	14.60	103.22	1322701.00	75.05	77.60	16.70	92.60	79.20	7952.80	733.30
50%	0.80	71.90	74.60	13.80	27147.00	1199745.50	29746.50	0.00	539976.70	29846.50	18.05	104.01	1848169.00	81.70	80.85	19.00	97.30	91.85	10943.01	1506.82
70%	1.50	77.72	76.60	15.70	30610.50	1942652.00	49952.50	1.00	881439.35	45752.50	22.45	105.78	3225757.00	86.35	83.90	21.70	99.70	99.60	16848.34	2377.82
90%	3.40	82.52	80.10	20.80	41847.50	3779787.00	97000.50	1.00	2008522.60	82152.00	30.15	107.60	6667487.50	90.15	89.20	27.15	100.00	100.00	24440.81	5574.30
max	10.40	100.00	90.40	37.80	87972.00	12666565.00	759324.00	1.00	24471160.40	353944.00	45.00	111.88	64926530.00	97.70	97.40	34.70	100.00	100.00	45687.84	22632.00

Рисунок Б.4 – Описательная статистика

```

years = 2021
year_df = ecom.loc[ecom['year']==years]

year1 = pd.Categorical(year_df.year)
year_df = year_df.set_index(["cod", "year"])

f1=pd.DataFrame(round(year_df.describe(percentiles=[0.1, 0.3,0.5, 0.7,0.9]),2))
#f1.to_excel('Описательная статистика '+str(years)+'.xlsx')

region = 'Южный федеральный округ'
reg_df = ecom.loc[ecom['region']==region]

year1 = pd.Categorical(reg_df.region)
reg_df = reg_df.set_index(["cod", "year"])

f2=pd.DataFrame(round(reg_df.describe(percentiles=[0.1, 0.3,0.5, 0.7,0.9]),2))
#f2.to_excel(f'Описательная статистика по региону {region}.xlsx')

```

Рисунок Б.5 – Описательная статистика по годам и по регионам

```

"Логорифмирование всех переменных"
ecom['ln_eshare']=ecom['eshare'].apply(np.log)
ecom['ln_citypop']=ecom['citypop'].apply(np.log)
ecom['ln_youthpop']=ecom['youthpop'].apply(np.log)
ecom['ln_poorpop']=ecom['poorpop'].apply(np.log)
ecom['ln_incomerpop']=ecom['incomerpop'].apply(np.log)
ecom['ln_pop']=ecom['pop'].apply(np.log)
ecom['ln_students']=ecom['students'].apply(np.log)
ecom['ln_grp']=ecom['grp'].apply(np.log)
ecom['ln_entrepreneur']=ecom['entrepreneur'].apply(np.log)
ecom['ln_investment']=ecom['investment'].apply(np.log)
ecom['ln_cpi']=ecom['cpi'].apply(np.log)
ecom['ln_card']=ecom['card'].apply(np.log)
ecom['ln_internetorg']=ecom['internetorg'].apply(np.log)
ecom['ln_internetpop']=ecom['internetpop'].apply(np.log)
ecom['ln_sctec']=ecom['sctec'].apply(np.log)
ecom['ln_citydig']=ecom['citydig'].apply(np.log)
ecom['ln_condig']=ecom['condig'].apply(np.log)
ecom['ln_roads']=ecom['roads'].apply(np.log)
ecom['ln_cargo']=ecom['cargo'].apply(np.log)

```

Рисунок Б.6 – Логарифмирование показателей и добавление их в датафрейм

```

plt.rcParams['figure.dpi'] = 100
sns.displot(data=ecom, x='eshare', kde=True, hue="covid", stat="percent")

```

Рисунок Б.6 – Создание графика распределения

```

sns.catplot(
    data=ecom, x="covid", y="eshare",
    kind="box"
)

```

Рисунок Б.7 – Создание графика «ящик с усами»

```
cormat=ecom[['eshare', 'citypop', 'youthpop', 'poorpop', 'incomepop', 'pop', 'students',
            'covid', 'grp', 'entrepreneur', 'investment', 'cpi', 'card', 'internetorg',
            'internetpop', 'sctec', 'citydig', 'condig', 'roads', 'cargo']].corr()
sns.heatmap(round(cormat,2), cmap='coolwarm', vmin=-1, vmax=1, annot=True)
```

Рисунок Б.8 – Создание корреляционной матрицы

```
]# определяем зависимую переменную и набор независимых переменных
x = elog[['citypop', 'youthpop', 'ln_incomepop', 'covid', 'ln_cpi', 'ln_internetorg', 'ln_internetpop', 'ln_cargo']]
y = elog['ln_eshare'].tolist()

# добавляем константу
x = sm.add_constant(x)

# Оцениваем модель регрессии используя МНК
result = sm.OLS(y, x).fit()

# Вывод таблицы с результатами оценки уравнения регрессии
print(result.summary())
```

Рисунок Б.9 – Построение модели МНК

```
vif = pd.DataFrame()
vif["feature"] = x.columns
vif["VIF"] = [variance_inflation_factor(x.values, i)
              for i in range(len(x.columns))]
print(vif)
```

Рисунок Б.10 – Расчет параметра VIF

```
#Влиятельные наблюдения
fig = sm.graphics.influence_plot(result, criterion="cooks")
fig.tight_layout(pad=1.0)
plt.axhline(y=2, color='r', linestyle='--')
plt.axhline(y=-2, color='r', linestyle='--')
plt.show()
fig = sm.graphics.plot_leverage_resid2(result)
fig.tight_layout(pad=1.0)
plt.show()
```

Рисунок Б.10 – Построение графика расстояний Кука и нормализованных остатков

```
pred_val = result.fittedvalues.copy()
true_val = elog['ln_eshare'].values.copy()
residual = true_val - pred_val
fig, ax = plt.subplots(figsize=(6,2.5))
plt.axhline(y=0, color='g', linestyle='--')
ax.scatter(pred_val, residual)
```

Рисунок Б.11 – Построение графика распределения остатков модели

```
#Тест Бреуша-Пагана
labels_bp = ['X square statistic', 'p-value',
            'f-value', 'f p-value']
test_bp = sm.stats.diagnostic.het_breuschpagan(result.resid, result.model.exog)

from statsmodels.compat import lzip
lzip(labels_bp, test_bp)

[('X square statistic', 29.477253890533603),
 ('p-value', 0.0002613925668331232),
 ('f-value', 3.874461350587111),
 ('f p-value', 0.0002018559307405745)]
```

Рисунок Б.12 – Проведение теста Бреуша-Пагана на определение гетероскедастичности модели

```
labels_g = ['Goldfeld Quandt statistic', 'p-value']
test_g = sm.stats.diagnostic.het_goldfeldquandt(result.resid, result.model.exog)

lzip(labels_g, test_g)

[('Goldfeld Quandt statistic', 1.5827655013986335),
 ('p-value', 0.000546065060668195)]
```

Рисунок Б.13 – Проведение теста Голдфелда-Квандта на определение гетероскедастичности модели

```
fig = plt.figure(figsize=(15, 30))
sm.graphics.plot_partregress_grid(result, fig=fig)
plt.show()
```

Рисунок Б.14 – Построение графиков частичной регрессии

```
fig = plt.figure(figsize=(15, 30))
sm.graphics.plot_ccpr_grid(result, fig=fig)
plt.show()
```

Рисунок Б.15 – Построение графиков частичных остатков

```
#Тесты на спецификацию
label_rainbow = ["Rainbow test statistic", "p - value"]
test_rainbow = sm.stats.diagnostic.linear_rainbow(result)
lzip(label_rainbow, test_rainbow)

[('Rainbow test statistic', 0.9025198646266417),
 ('p - value', 0.7708126944499328)]
```

Рисунок Б.16 – Проведение Linear Rainbow Test на определение спецификации модели

```
label_reset = ["Wald ", "p - value"]
test_reset = sm.stats.diagnostic.linear_reset(result, power=5, test_type = "fitted")
test_reset
```

```
/usr/local/lib/python3.10/dist-packages/statsmodels/stats/diagnostic.py:1081: FutureWarning
    aug = res.fittedvalues[:, None]
<class 'statsmodels.stats.contrast.ContrastResults'>
<Wald test (chi2): statistic=8.877219984025986, p-value=0.06424274499366901, df_denom=4>
```

Рисунок Б.17 – Проведение RESET-тест Рэмзи на определение спецификации модели

```
#Нормальность распределения
import scipy as sp
fig, ax = plt.subplots(figsize=(6,2.5))
sp.stats.probplot(residual, plot=ax, fit=True)
sns.histplot(residual,kde=True, stat="percent")
```

Рисунок Б.18 – Проверка на нормальность распределения остатков (построение Q-Q графика)

```
x = elog[['citypop', 'youthpop', 'ln_incomepop', 'covid', 'ln_entrepreneur',
        'ln_cpi', 'ln_internetorg', 'ln_internetpop', 'ln_sctec', 'ln_cargo']]
y = elog['ln_eshare'].tolist()

# добавляем константу
x = sm.add_constant(x)

# Оцениваем модель регрессии используя МНК
model_OLS = PooledOLS(elog.ln_eshare, x)
result_OLS = model_OLS.fit()
# Вывод таблицы с результатами оценки уравнения регрессии
print(result_OLS)
```

Рисунок Б.19 – Построение модели PooledOLS

```
model_RE = RandomEffects(elog.ln_eshare, x)
result_RE = model_RE.fit()
print(result_RE)
```

Рисунок Б.20 – Построение модели RandomEffects

```
#фиксированные эффекты, рассчитанные с использованием within transformation
model_FE_i = PanelOLS(elog2.ln_eshare, x, entity_effects=True)
result_FE_i = model_FE_i.fit()
print(result_FE_i)
```

Рисунок Б.21 – Построение модели Fixed Effects

## ПРИЛОЖЕНИЕ В.

### Промежуточные результаты расчетов и проверок на гипотезы

Таблица В.1 – Выделенные показатели из литературного обзора

Показатель	Источник данных	Период данных	Охват данных	Тип переменной	Влияние	Значимость	Авторы
Индекс электронной коммерции В2С	Доклада об информационной экономике ЮНКТАД	2014	149 стран мира	у			A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat
Оценка субиндекса готовности (инфраструктура ИКТ)	Глобальный отчет по информационным технологиям, опубликованным Всемирным экономическим форумом	2014	149 стран мира	тех х	+	высокая значимость	A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat
Количество подписок на мобильный интернет	Глобальный отчет по информационным технологиям, опубликованным Всемирным экономическим форумом	2014	149 стран мира	тех х	+	низкая значимый	A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat
Использование соц сетей	Глобальный отчет по информационным технологиям, опубликованным Всемирным экономическим форумом	2014	149 стран мира	тех х	+	средняя значимость	A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat
Доля городского населения, в процентах от общей численности населения	Сайт Всемирного банка	2014	149 стран мира	соц х	+	средняя значимость	A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat
Ожидаемая продолжительность обучения, в годах	Доклад о человеческом развитии	2014	149 стран мира	соц х	+	средняя значимость	A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat

Продолжение таблицы В.1

Индекс кибербезопасности	Глобальный индекс кибербезопасности (GCI), разработанный ABI Research	2014	149 стран мира	тех х	+	низкая значимость	A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat
ВНД на душу населения, в процентах	Доклад о человеческом развитии	2014	149 стран мира	эк х	+	очень высокая значимость	A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat
Индекс эффективности логистики	Отчет «Ведение бизнеса», опубликованного Всемирным банком	2014	149 стран мира	эк х	+	средняя значимость	A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat
Расстояние до границы	Отчет «Ведение бизнеса», опубликованного Всемирным банком	2014	149 стран мира	эк х	+	низкая значимость	A. Waseem, Yasir Rashid, M. Warraich, I. Sadiq, Zeeshan Shaukat
Объем интернет-покупок физических лиц, в процентах от общего числа физических лиц		2015	31 европейских стран	у			K. Dumičić, B. Žmuk, B. Novkovska
ВВП на душу населения, в процентах	Евростат	2015	31 европейских стран	эк х	+	высокая значимость	K. Dumičić, B. Žmuk, B. Novkovska
Государственные расходы на образование, в процентах от ВВП	Евростат, Всемирный Банк	2015	31 европейских стран	эк х	+	высокая значимость	K. Dumičić, B. Žmuk, B. Novkovska
Доля лиц, пользующихся Интернетом, в процентах	Евростат	2015	31 европейских стран	тех х	+	очень высокая значимость	K. Dumičić, B. Žmuk, B. Novkovska
Доля лиц в возрасте от 16 до 74 лет, которые смогли выполнить 1 или 2 из 6 видов деятельности, связанных с Интернетом, в процентах	Евростат	2015	31 европейских стран	тех х	+	средняя значимость	K. Dumičić, B. Žmuk, B. Novkovska
Общий товарооборот компаний в сфере интернет-торговли, в процентном отношении	Евростат	2003-2017	13 европейских стран	у			Lucero Ortiz A., Rodríguez J. C., Gómez M

# Продолжение таблицы В.1

Количество подписок на мобильный Интернет на 100 человек	Евростат	2003-2017	13 европейских стран	тех х	+	высокая значимость	Lucero Ortiz A., Rodríguez J. C., Gómez M
Процент от ВВП, посвященный исследованиям и расходы на развитие	Евростат	2003-2017	13 европейских стран	эк х	+	средняя значимость	Lucero Ortiz A., Rodríguez J. C., Gómez M
Располагаемый доход на душу населения	Евростат	2003-2017	13 европейских стран	соц х	+	очень высокая значимость	Lucero Ortiz A., Rodríguez J. C., Gómez M
Объема продаж в сфере электронной коммерции B2C	Опрос AECEM и Red.es Обсерватория	1996-2003	Испания	у			I. Rodríguez-Ardura, A. Meseguer-Artola, J. Vilaseca-Requena
Количество пользователей Интернета	Отчет Испанской ассоциации исследований средств массовой информации	1996-2003	Испания	тех х	+	очень высокая значимость	I. Rodríguez-Ardura, A. Meseguer-Artola, J. Vilaseca-Requena
Распространенность широкополосной связи		1996-2003	Испания	тех х	+	высокая значимость	I. Rodríguez-Ardura, A. Meseguer-Artola, J. Vilaseca-Requena
Правовые рамки защиты прав потребителей в Интернете (появление закона)		1996-2003	Испания	соц х	+	средняя значимость	I. Rodríguez-Ardura, A. Meseguer-Artola, J. Vilaseca-Requena
Среднерыночная стоимость на самый производительный компьютер в евро	данные Европейской обсерваторией информационных технологий	1996-2003	Испания	эк х	+	средняя значимость	I. Rodríguez-Ardura, A. Meseguer-Artola, J. Vilaseca-Requena
Количества серверов с доменом ".es" - кодом страны для Испании, использующий протоколы безопасности на 1000 пользователей	данным, подготовленным для Испании сервисной фирмой Netcraft	1996-2003	Испания	тех х	+	средняя значимость	I. Rodríguez-Ardura, A. Meseguer-Artola, J. Vilaseca-Requena
Фиктивный значение - пользователь электронной коммерции	Национальный статистический институт Испании	2014 и 2019	Испания	у			Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega



Продолжение таблицы В.1

Использование компьютера	Национальный статистический институт Испании	2014 и 2019	Испания	тех х	+	очень высокая значимость	Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega
Использование Интернета	Национальный статистический институт Испании	2014 и 2019	Испания	тех х	+	очень высокая значимость	Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega
Пол	Национальный статистический институт Испании	2014 и 2019	Испания	соц х	+	средняя значимость	Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega
Возраст	Национальный статистический институт Испании	2014 и 2019	Испания	соц х	+	средняя значимость	Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega
Гражданство	Национальный статистический институт Испании	2014 и 2019	Испания	соц х	+	средняя значимость	Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega
Уровень образования	Национальный статистический институт Испании	2014 и 2019	Испания	соц х	+	средняя значимость	Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega
Уровень дохода	Национальный статистический институт Испании	2014 и 2019	Испания	соц х	+	средняя значимость	Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega
Уровень безопасности	Национальный статистический институт Испании	2014 и 2019	Испания	тех х	+	высокая значимость	Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega
Подключение к широкополосному Интернету	Национальный статистический институт Испании	2014 и 2019	Испания	тех х	+	очень высокая значимость	Fernando Fernández-Bonilla, Covadonga Gijón, Bárbara De la Vega
Объем розничных продаж через Интернет	Индиястат, Statista, KPMG, Digital Intelligence Special Report, cmocouncil, go-globe	2005-2020	Индия	у			Dr.Debesh Bhowmik
Количество пользователей Интернетом	Индиястат, Statista, KPMG, Digital Intelligence Special Report, cmocouncil, go-globe	2005-2020	Индия	тех х	+	очень высокая значимость	Dr.Debesh Bhowmik

Продолжение таблицы В.1

Количество пользователей кредитными картами	Индиястат, Statista, KPMG, Digital Intelligence Special Report, cmocouncil, go-globe	2005-2020	Индия	эк х	+	средняя значимость	Dr.Debesh Bhowmik
Количество пользователей дебетовыми картами	Индиястат, Statista, KPMG, Digital Intelligence Special Report, cmocouncil, go-globe	2005-2020	Индия	эк х	+	средняя значимость	Dr.Debesh Bhowmik
Темпы роста ВВП	Индиястат, Statista, KPMG, Digital Intelligence Special Report, cmocouncil, go-globe	2005-2020	Индия	эк х	+	высокая значимость	Dr.Debesh Bhowmik
Объем операций электронной коммерции (кол-во транзакций)	сайт ВКМ	2010-2014	Турция	у			Filiz Kadı, Canan Peker
Индекс потребительских цен по отношению к 2003 году	базы данных Центрального банка Турции	2010-2014	Турция	эк х	-	высокая значимость	Filiz Kadı, Canan Peker
Общее количество пользователей Интернета	сайт IAB Turkey	2010-2014	Турция	тех х	+	средняя значимость	Filiz Kadı, Canan Peker
Объем импорта в млн. tl	базы данных Центрального банка Турции	2010-2014	Турция	эк х	+	высокая значимость	Filiz Kadı, Canan Peker
Объем операций с кредитными картами (кол-во транзакций)	данные Card Monitor 2014, опубликованным ВКМ	2010-2014	Турция	эк х	+	высокая значимость	Filiz Kadı, Canan Peker
Доля рынка электронной коммерции в обороте розничной торговли	Росстат	2011-2017	Россия	у			Мотина В. Г.
Доля домашних хозяйств, имеющих доступ к сети интернет, в общем числе домашних хозяйств	Росстат	2011-2017	Россия	тех х	+	высокая значимость	Мотина В. Г.

# Продолжение таблицы В.1

Доля домашних хозяйств, использующих для выхода в сеть Интернет мобильные телефоны или смартфоны	Росстат	2011-2017	Россия	тех х	+	высокая значимость	Мотина В. Г.
Реальные располагаемые денежные доходы по РФ	Росстат	2011-2017	Россия	соц х	-	средняя значимость	Мотина В. Г.
Объем розничной торговли в сети интернет (млрд. руб.)	Росстат	2014	Россия	у			Видякин, А. Н.
Проникновение интернета (%)	Росстат	2014	Россия	тех х	+	средняя значимость	Видякин, А. Н.
Индекс доступности интернета	Росстат	2014	Россия	тех х	+	очень высокая значимость	Видякин, А. Н.
Коэффициент официальной ликвидации организаций	Росстат	2014	Россия	эк х	+	высокая значимость	Видякин, А. Н.
Доля автомобильных дорог общего пользования, не отвечающих нормативным требованиям (%)	Росстат	2014	Россия	тех х	+	низкая значимый	Видякин, А. Н.
Численность населения с доходами ниже прожиточного минимума (%)	Росстат	2014	Россия	соц х	+	средняя значимость	Видякин, А. Н.
Объем рынка электронной торговли, млрд.руб.	Росстат	2008-2018	Россия	у			Михайленко А.А.
Оборот розничной торговли, трлн.руб	Росстат	2008-2018	Россия	эк х	-	очень высокая значимость	Михайленко А.А.
Индексы потребительский цен на товары и услуги, %	Росстат	2008-2018	Россия	эк х	не проверялась	не проверялась	Михайленко А.А.
Среднемесячная номинальная заработная плата, руб.	Росстат	2008-2018	Россия	соц х	не проверялась	не проверялась	Михайленко А.А.
Среднедушевые денежные доходы населения, руб.	Росстат	2008-2018	Россия	соц х	+	очень высокая значимость	Михайленко А.А.
Индекс потребительской уверенности, %	Росстат	2008-2018	Россия	соц х	не проверялась	не проверялась	Михайленко А.А.

Продолжение таблицы В.1

Доля онлайн-продаж в розничных продажах	Федеральная служба государственной статистики РФ; Ассоциация компаний интернет-торговли (АКИТ); ПАО «Сбербанк»	2011-2021	Россия	у			Коротких В.В.
Проникновение онлайн-торговли в России, оцененное как доля городских жителей в возрасте 16-55 лет, покупающие онлайн чаще раза в год	Федеральная служба государственной статистики РФ; Международный союз электросвязи (International Telecommunication Union)	2011-2021	Россия	тех х	+	высокая значимость	Коротких В.В.
Индекс эффективности логистики	Всемирный банк (WorldBank); Всемирный почтовый союз (Universal Postal Union)	2011-2021	Россия	тех х	+	высокая значимость	Коротких В.В.
Индекс человеческого капитала	Отдел государственного управления и управления развитием департамента по экономическим и социальным вопросам ООН	2011-2021	Россия	соц х	+	высокая значимость	Коротких В.В.
Индекс качества онлайн-обслуживания	Отдел государственного управления и управления развитием департамента по экономическим и социальным вопросам ООН	2011-2021	Россия	соц х	+	низкая значимый	Коротких В.В.

## Окончание таблицы В.1

Доля онлайн-заказов, получаемых в пункте выдачи	Яндекс.Маркет; ООО «ГфК-Русь» (Growth from Knowledge Rus)	2011-2021	Россия	тех х	+	средняя значимость	Коротких В.В.
Переменная для учета фиксированных эффектов в динамике электронной коммерции, связанных с COVID-19		2011-2021	Россия	соц х	+	высокая значимость	Коротких В.В.

Таблица В.2 – Описание выделенных для анализа показателей

Показатель	Ед.изм	Наименование	Тип переменной	Источник
eshare	%	Доля продаж через Интернет в общем объеме оборота розничной торговли, в фактически действовавших ценах	у	Росстат
citypop	%	Доля городского населения, от общей численности населения	соц х	Росстат
youthpop	%	Доля населения в трудоспособном возрасте, от общей численности населения	соц х	Росстат
poorpop	%	Численность населения с денежными доходами ниже границы бедности (величины прожиточного минимума), от общей численности населения	соц х	Росстат
incomerpop	руб/мес.	Среднедушевые денежные доходы населения руб./мес.	соц х	Росстат
pop	человек	Численность постоянного населения в среднем за год	соц х	Росстат
students	человек	Численности студентов образовательных организаций, осуществляющих образовательную деятельность по образовательным программам высшего образования	соц х	Министерство образования
covid		Фиктивная переменная: до 2020 - 0, после 2020 - 1	соц х	
grp	млн.руб	Валовой региональный продукт в тек.ценах	эк х	Росстат

## Окончание таблицы В.2

entrepreneur	человек	Численность индивидуальных предпринимателей в среднем за год, сведения о которых содержатся в Едином государственном реестре индивидуальных предпринимателей	эк х	ФНС
investment	%	Доля инвестиций, направленных на реконструкцию и модернизацию, в общем объеме инвестиций в основной капитал	эк х	Росстат
cpi	индекс	Индекс потребительских цен, к прошлому году	эк х	Росстат
card	шт	Количество платежных карт, выданных физ лицам на территории региона на конец периода	тех х	Банк России
internetorg	%	Доля организаций, использующих широкополосный доступ к сети Интернет, в общем числе организаций (значение показателя за год)	тех х	Росстат
internetpop	%	Использование сети Интернет населением в возрасте 15 лет и старше, от общей численности населения в возрасте 15 лет и старше, соответствующего субъекта Российской Федерации	тех х	Росстат
sctec	%	Доля продукции высокотехнологичных и наукоемких отраслей в валовом региональном продукте	тех х	Росстат
citydig	%	Уровень цифровизации местной телефонной сети в городской местности по субъектам Российской Федерации	тех х	Минцифры России
condig	%	Уровень цифровизации местной телефонной сети в сельской местности по субъектам Российской Федерации	тех х	Минцифры России
roads	км	Протяженность автомобильных дорог общего пользования федерального, регионального или межмуниципального и местного значения с твердым покрытием на конец года	лог х	Росстат
cargo	млн.тонно-км	Грузооборот автомобильного транспорта по субъектам Российской Федерации	лог х	Росстат

Таблица В.3 – Описательная статистика для всех данных

Показатель	Кол-во наблюдений	Среднее значение	Стандартное отклонение	Минимальное значение	Квартили распределения					Максимальное значение
	count	mean	std	min	10%	30%	50%	70%	90%	max
esales	486	10814,42	41757,65	0	125,74	618,17	1679,27	4783,19	16972,56	555569,3
eshare	486	1,38	1,66	0	0,1	0,4	0,8	1,5	3,4	10,4
citypop	486	70,39	13,09	29	53,1	66,25	71,9	77,72	82,52	100
youthpop	486	75,22	4,26	68,7	70,9	72,5	74,6	76,6	80,1	90,4
poorpop	486	14,52	5,04	5	8,85	12,2	13,8	15,7	20,8	37,8
incomepop	486	29654,6	10182,19	14963	20631,5	24156,5	27147	30610,5	41847,5	87972
pop	486	1808925	1820073	138401	465844	865581,5	1199746	1942652	3779787	12666565
students	486	51081,44	86937,63	1947	7866,5	18629,5	29746,5	49952,5	97000,5	759324
covid	486	0,33	0,47	0	0	0	0	1	1	1
grp	486	1140211	2413590	47434,9	153425,3	315598,5	539976,7	881439,4	2008523	24471160
entrepreneur	486	44736,52	51276,79	2434	11587	18578,5	29846,5	45752,5	82152	353944
investment	486	18,85	7,79	3,5	8,9	14,6	18,05	22,45	30,15	45
cpi	486	104,58	2,02	100,84	102,39	103,22	104,01	105,78	107,6	111,88
card	486	3469287	6413238	184566	539955,5	1322701	1848169	3225757	6667488	64926530
internetorg	486	78,72	11,4	29	61,2	75,05	81,7	86,35	90,15	97,7
internetpop	486	81,11	5,86	65,3	73,75	77,6	80,85	83,9	89,2	97,4
setec	486	19,32	5,53	6,7	11,95	16,7	19	21,7	27,15	34,7
citydig	486	93,86	8,49	44,7	82,55	92,6	97,3	99,7	100	100
condig	486	85,66	16,51	22,4	61,5	79,2	91,85	99,6	100	100
roads	486	13340,66	8752,6	855,86	3539,36	7952,8	10943,01	16848,34	24440,81	45687,84
cargo	486	2253,19	2561,51	9,9	185,46	733,3	1506,82	2377,82	5574,3	22632

Таблица В.4 – Описательная статистика для данных за 2016 год

Показатель	Кол-во наблюдений	Среднее значение	Стандартное отклонение	Минимальное значение	Квартили распределения					Максимальное значение
	count	mean	std	min	10%	30%	50%	70%	90%	max
esales	81	4313,06	15922,25	0	0	153,94	597,34	1382,63	8485,96	132754,7
eshare	81	0,56	0,8	0	0	0,1	0,3	0,5	1,4	4,6
citypop	81	69,93	13,47	29,2	52,2	65,7	71,7	77,2	82,1	100
youthpop	81	75,92	4,32	70,1	71,8	73,1	75,1	77,3	80,8	90,4
poorpop	81	15,54	5,27	7,5	9,8	13	14,7	16,8	22,1	37,8
incomepop	81	26693,01	8228,28	14963	18859	22510	25081	27849	37359	62004
pop	81	1810179	1801506	145958	467115	863354	1223135	1975574	3637758	12355395
students	81	53976,44	88544,83	2309	8948	19124	32078	52127	102754	736274
grp	81	914172,9	1862925	47434,9	148810,1	279140,7	431549,8	762510,3	1421518	15121534
entrepreneur	81	42910,4	44134,44	3349	12831	19639	29762	45436	79827	244323
investment	81	19,04	7,47	4,3	8,7	14,8	18,3	23,3	27,9	34,2
cpi	81	107,35	1,12	104,61	106,07	106,68	107,38	107,87	108,53	111,88
card	81	3124954	6064780	184566	423305	1194417	1654657	3107911	5972418	52692706
internetorg	81	81,77	7,27	56,2	72,2	78,7	82,9	85,8	90	95,5
internetpop	81	78,66	4,89	67,9	72,9	76,3	78,6	81,4	84,2	91,3
sctec	81	18,31	4,94	7,4	11,3	15,9	17,8	20,6	24,5	31,2
citydig	81	91,07	9,58	44,7	80,3	87,5	93,3	97,4	100	100
condig	81	81,31	18,76	22,4	56,5	72,9	87,3	96,4	100	100
roads	81	13000,24	8564,83	855,86	3504,28	7841,96	10794,63	16696,89	23339	42989,42
cargo	81	1885,34	1972,51	9,9	221,26	683,22	1393,63	2219,57	3781,85	11176,3



Таблица В.4 – Описательная статистика для данных за 2021 год

Показатель	Кол-во наблюдений	Среднее значение	Стандартное отклонение	Минимальное значение	Квартили распределения					Максимальное значение
	count	mean	std	min	10%	30%	50%	70%	90%	max
esales	81	24759,04	71541,7	89,84	657,57	4256,43	6511,14	13434,33	46675,73	555569,3
eshare	81	3,11	2,24	0,1	0,8	1,7	2,4	3,8	5,6	10,4
citypop	81	70,7	13,04	29,1	54,9	66,7	72,2	77,9	82,7	100
youthpop	81	75,1	4,16	69,3	70,9	72,4	74,7	76,7	79,8	89,5
poorpop	81	13,35	4,64	5	7,9	11,4	12,8	14,4	19,3	29,3
incomepop	81	33884,56	11780,08	17954	23773	27788	30505	34918	49938	87972
pop	81	1800179	1843984	138401	484898	869839	1175726	1902148	3792279	12645258
students	81	49679,94	91160,37	1957	7571	18194	27747	46567	95712	759324
grp	81	1494405	3142331	71336,2	199326,3	421672,8	690253,2	1083779	2122537	24471160
entrepreneur	81	43432,89	55048,99	2434	10924	17032	27559	41303	77718	353944
investment	81	18,66	7,77	4,4	8,9	14,7	18	21,2	28	45
cpi	81	106,79	1	104,32	105,49	106,35	106,86	107,41	107,93	108,96
card	81	4062006	7570133	259194	692187	1644919	2236208	3825353	7730676	64926530
internetorg	81	76,47	5,2	57,5	70,6	73,9	76	79,4	82,2	87,6
internetpop	81	85,97	4,54	76,3	80,6	83,1	85,1	88	92,2	97,4
sctec	81	19,42	6,21	8,2	10,4	16,2	18,9	22	29,2	33,5
citydig	81	96,2	6,49	66,2	87,2	96,9	99,2	100	100	100
condig	81	90,17	13,33	42	73,4	88,4	96,5	100	100	100
roads	81	13660,11	8974,87	1071,11	3654,14	8271,53	11100,82	17203,09	24838,99	45687,84
cargo	81	2651,94	3207,46	21,7	114,29	778,52	1597,63	2381,16	6873,17	14074,42

Таблица В.5 – Результаты моделирования данных (создание множественной регрессионной модели)

Переменная	Тип и номер модели					
	y-x	y-lnx	lny-x	lny-lnx	lny-lnx	lny-lnx
	eshare	eshare	ln_eshare	ln_eshare	ln_eshare	ln_eshare
	1	2	3	4	5	6
const	-6,0732** (2,964)	-46,1857*** (14,430)	-0,4939 (2,171)	-10,3485 (10,173)	-15,6215** (5,760)	-2,8669 (8,745)
citypop	0,0180*** (0,006)		0,0119** (0,004)			
youthpop	-0,0062 (0,021)		-0,0149 (0,018)			
poorpop	-0,0191 (0,019)		0,0033 (0,017)			
incomepop	-1,599e-05 (1,06e-05)		6,868e-07 (1e-05)		1,409e-06 (1,02e-05)	
students	4,003e-07 (2,42e-06)		2,929e-06* (1,6e-06)		2,821e-06* (1,59e-06)	
covid	1,9553*** (0,182)	2,0076*** (0,195)	1,5857*** (0,143)	1,3584*** (0,139)	1,5464*** (0,140)	1,5797*** (0,113)
grp	-1,232e-07 (8,97e-08)		-7,377e-08 (6,04e-08)		-4,352e-08 (6,06e-8)	
entrepreneur	9,982e-06*** (2,82e-06)		5,385e-06** (1,86e-06)		4,889e-06** (1,85e-06)	
investment	-0,0184** (0,007)		-0,0085 (0,006)			
cpi	0,0431* (0,025)		-0,0436** (0,018)		-0,0452** (0,018)	
card	5,167e-08 (5,81e-08)		-2,525e-08 (3,81e-08)		-3,167e-08 (3,79e-08)	
internetorg	0,0100 (0,007)		0,0158** (0,005)			
internetpop	0,0076 (0,013)		0,0244** (0,009)			
sctec	0,0156 (0,011)		0,0216** (0,007)			

Продолжение таблицы В.5

citydig	0,0054 (0,007)		-0,0123** (0,005)			
condig	-0,0006 (0,004)		-0,0038 (0,003)			
roads	-3,215e- 05*** (8,43e-06)		-5,686e-06 (5,76e- 06)		-5,828e-06 (5,61e-06)	
cargo	0,0002*** (3,41e-05)		5,159e-05** (2,24e- 05)		4,681e-05 (2,26e-05)	
ln_citypop		0,1132 (0,394)		-0,1437 (0,298)	0,4705** (0,260)	1,5446*** (0,295)
ln_youthpop		1,5548 (1,764)		0,643 (1,274)	-0,4184 (1,230)	-4,4520*** (1,023)
ln_poorpop		-0,2728 (0,329)		-0,0279 (0,235)	-0,1739 (0,231)	
ln_incomepop		0,7422* (0,415)		0,8203** (0,318)		0,7527*** (0,203)
ln_students		0,0453 (0,139)		-0,1875* (0,105)		0,1140** (0,052)
ln_grp		-0,8572** (0,269)		-0,5195** (0,184)		
ln_entrepreneur		0,7954*** (0,242)		-0,0288 (0,180)		
ln_investment		-0,2992** (0,129)		-0,0355 (0,107)	-0,1353 (0,107)	
ln_cpi		4,2675 (2,781)		-4,264** (1,922)		-4,0189** (1,636)
ln_card		0,8224** (0,340)		0,877*** (0,269)		
ln_internetorg		1,0980** (0,482)		1,0282** (0,381)	1,1600** (0,385)	1,6705*** (0,334)
ln_internetpop		0,4816 (1,059)		1,8784** (0,749)	2,1295** (0,741)	1,4829** (0,625)
ln_sctec		0,3427 (0,224)		0,6021*** (0,148)	0,5805*** (0,137)	0,2890* (0,120)
ln_citydig		0,2705 (0,695)		0,3700 (0,472)	1,2102** (1,2102)	1,2005** (0,394)
ln_condig		0,1957 (0,299)		-0,2299 (0,191)	-0,3489* (0,190)	
ln_roads		-0,8071*** (0,139)		-0,173 (0,0,97)		0,4028*** (0,076)
ln_cargo		0,2906*** (0,082)		0,2722*** (0,066)		

Окончание таблицы В.5

Q	486	486	408	408	408	330
R-squared	0,599	0,581	0,610	0,626	0,615	0,743
F-statistic	38,76	35,99	33,78	36,24	34,48	76,22
Prob (F-statistic)	9,33e-81	1,97e-76	3,79e-68	9,9e-72	3,53e-69	7,9e-86
AIC	1465	1486	871,7	854	866,6	507
BIC	1545	1566	947,9	930,2	942,8	556,4
Omnibus	231,145	181,111	2,296	5,424	3,115	5,974
Prob(Omnibus)	0	0	0,317	0,066	0,211	0,050
Durbin-Watson	0,956	0,900	0,803	0,795	0,815	0,945
Jarque-Bera	1543,948	795,300	2,268	6,778	3,448	6,09
Prob(JB)	0	6,56e-167	0,322	0,0337	0,178	0,0474
*** p<0.01, ** p<0.05, * p<0.1						

Таблица В.6 – Рассчитанные параметры VIF для всех исследуемых моделей

Переменная	Тип и номер модели					
	y-x	y-lnx	lny-x	lny-lnx	lny-lnx	lny-lnx
	eshare	eshare	ln_eshare	ln_eshare	ln_eshare	ln_eshare
	1	2	3	4	5	6
const	3716,498	84341,99	4056,891	93060,06	28923,11	96389,24
citypop	2,2895		2,073327			
youthpop	3,481766		2,906658			
poorpop	4,025582		4,856984			
incomepop	4,936113		7,490378		7,913098	
students	18,73472		19,0915		19,15666	
covid	3,105097	3,421035	3,91862	3,834159	3,788003	3,561562
grp	19,78853		21,16537		21,62432	
entrepreneur	8,840656		8,685474		8,662324	
investment	1,271411		1,358126			
cpi	1,113617		1,178257		1,208919	
card	58,5545		59,09253		59,25718	
internetorg	2,321713		2,865936			
internetpop	2,326668		2,420684			
sctec	1,458056		1,470712			
citydig	1,701871		1,708958			
condig	1,746834		1,685661			
roads	2,298183		2,165135		2,08368	
cargo	3,230277		3,018176		3,09378	
ln_citypop		2,771023		2,597001	1,920363	1,485495
ln_youthpop		3,759880		2,714879	2,454485	1,524843
ln_poorpop		4,913281		4,83161	4,562026	
ln_incomepop		5,861071		6,430507		2,292999
ln_students		8,426110		9,335373		2,213367
ln_grp		35,544266		32,18445		
ln_entrepreneur		19,116052		18,48583		
ln_investment		1,481693		1,584558	1,528378	1,39645

Окончание таблицы В.6

ln_cpi		1,154176		1,223082		1,194307
ln_card		45,575870		50,12202		
ln_internetorg		2,448233		2,830631	2,804287	2,702454
ln_internetpop		2,377782		2,452954	2,329506	1,959176
ln_sctec		1,893025		1,913344	1,583003	1,63974
ln_citydig		2,125861		2,346942	1,952591	1,111784
ln_condig		2,055724		2,035875	1,952845	
ln_roads		4,280717		3,284427		2,230128
ln_cargo		4,982192		4,168348		

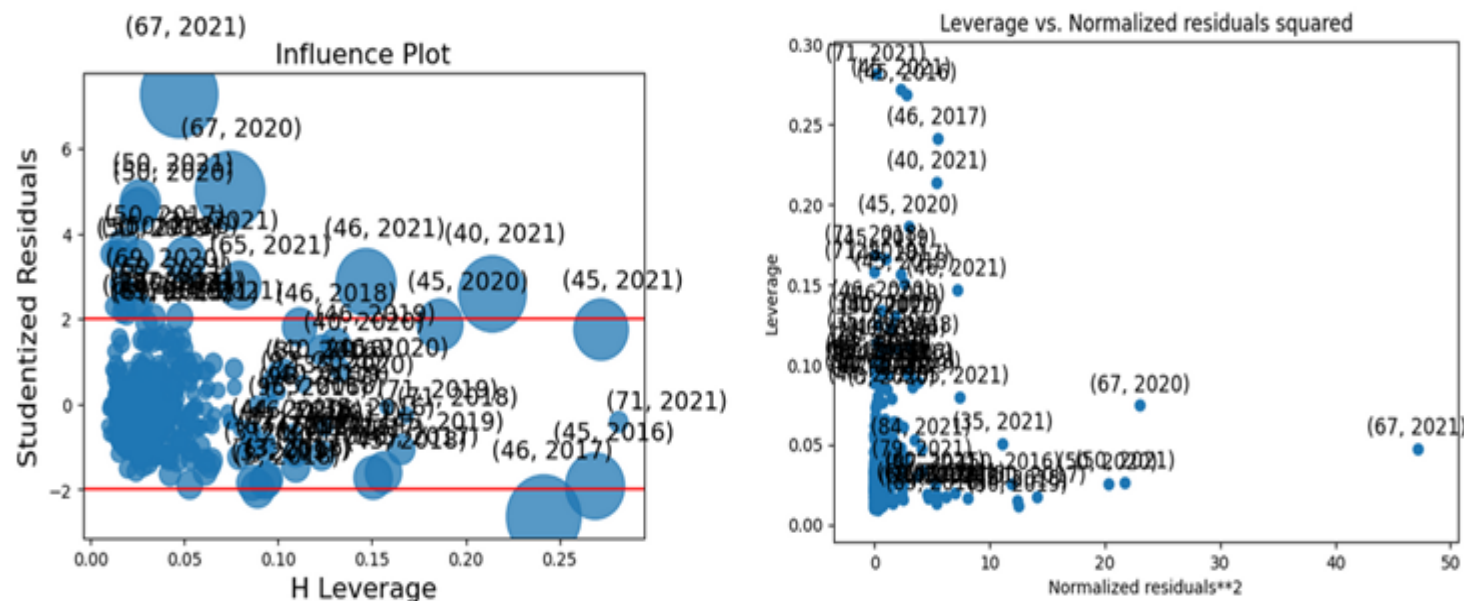


Рисунок В.1 – Графики расстояния Кука и нормализованных остатков для линейной модели №1

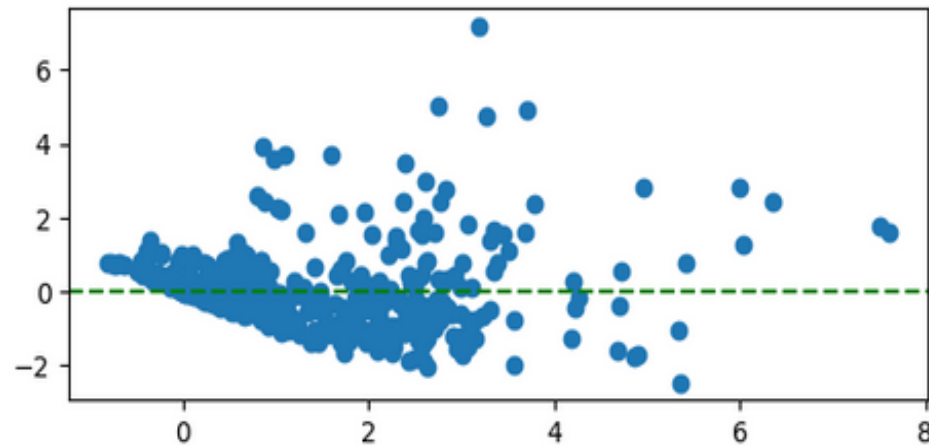


Рисунок В.2 – График распределения остатков по переменной eshare для модели №1

```
[('X square statistic', 76.92124843256939),
 ('p-value', 2.945734403508282e-09),
 ('f-value', 4.980445942515512),
 ('f p-value', 2.999611165514401e-10)]
```

Рисунок В.3 – Результаты теста Бреуша-Пагана для модели №1

```
[('Goldfeld Quandt statistic', 1.2365930867288217),
 ('p-value', 0.0697323609795023)]
```

Рисунок В.4 – Результаты теста Голдфелда-Квандта для модели №1

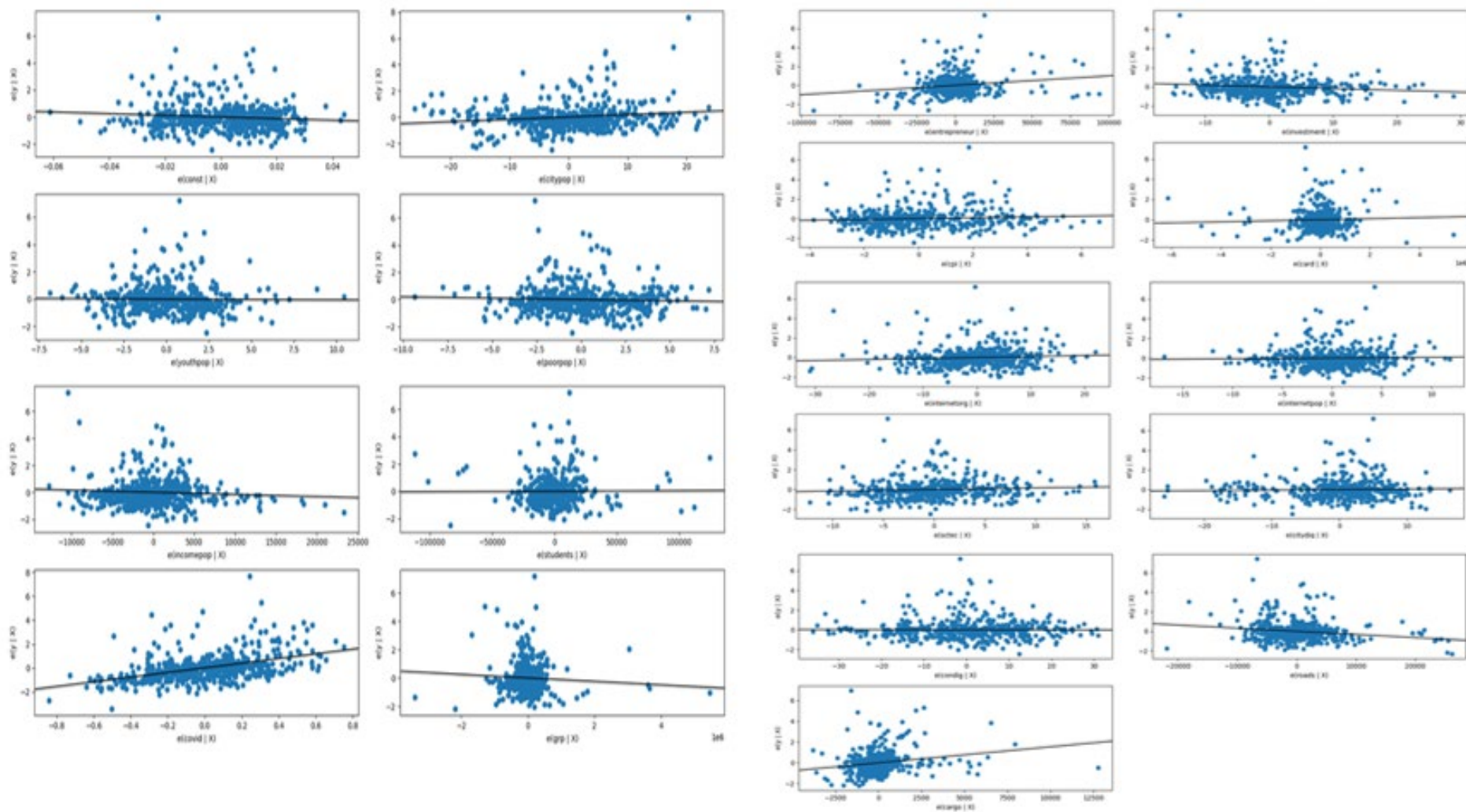


Рисунок В.5 – Графики частичной регрессии для модели №1



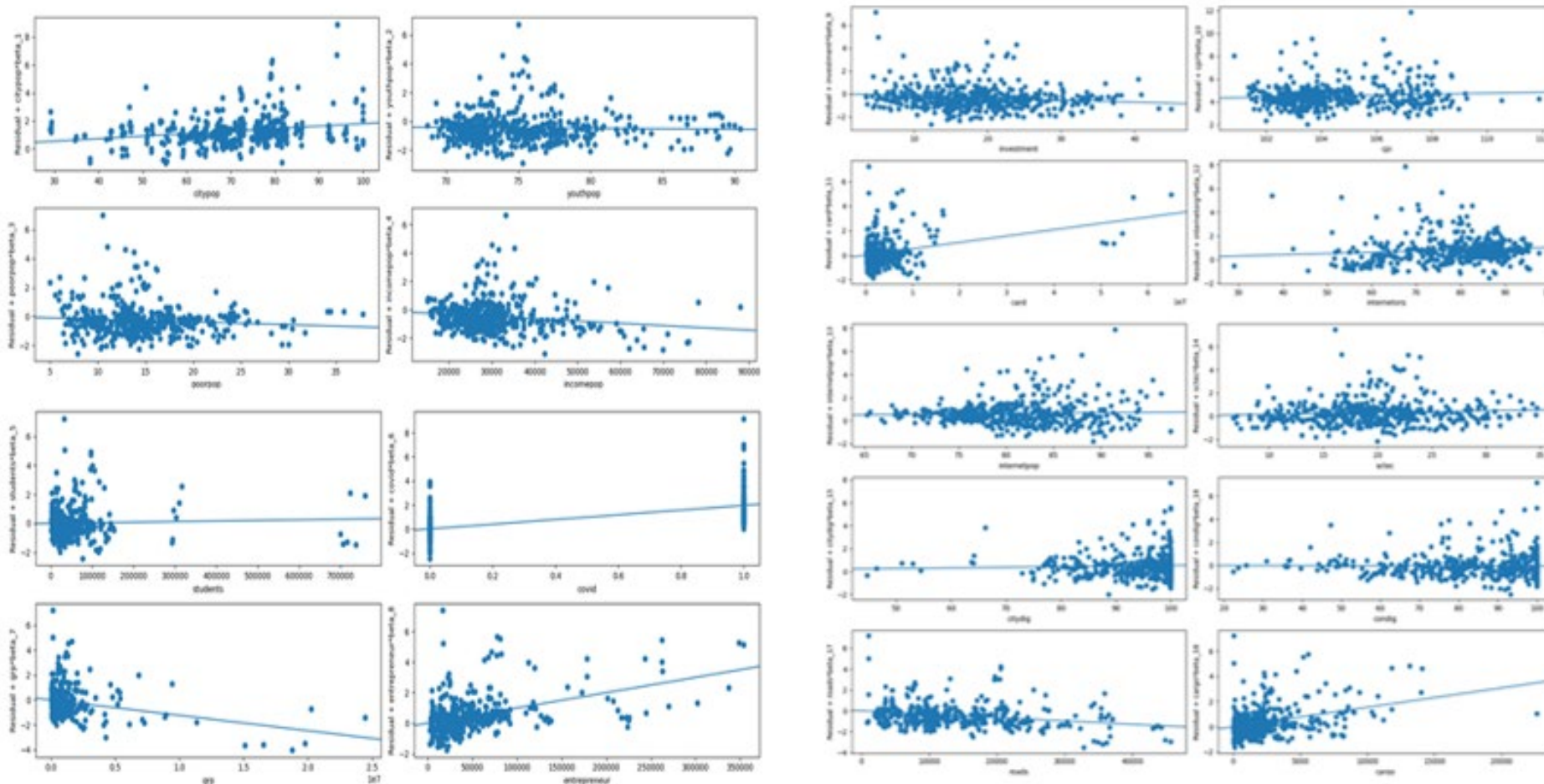


Рисунок В.6 – Графики частичных остатков для модели №1

```
[('Rainbow test statistic', 0.6693573396007154),
 ('p - value', 0.9978949614918331)]
```

Рисунок В.7 – Результаты Linear Rainbow Test для модели №1

```
<Wald test (chi2): statistic=97.33210723322614, p-value=3.6363117671473156e-20, df_denom=4>
```

Рисунок В.8 – Результаты RESET-теста для модели №1

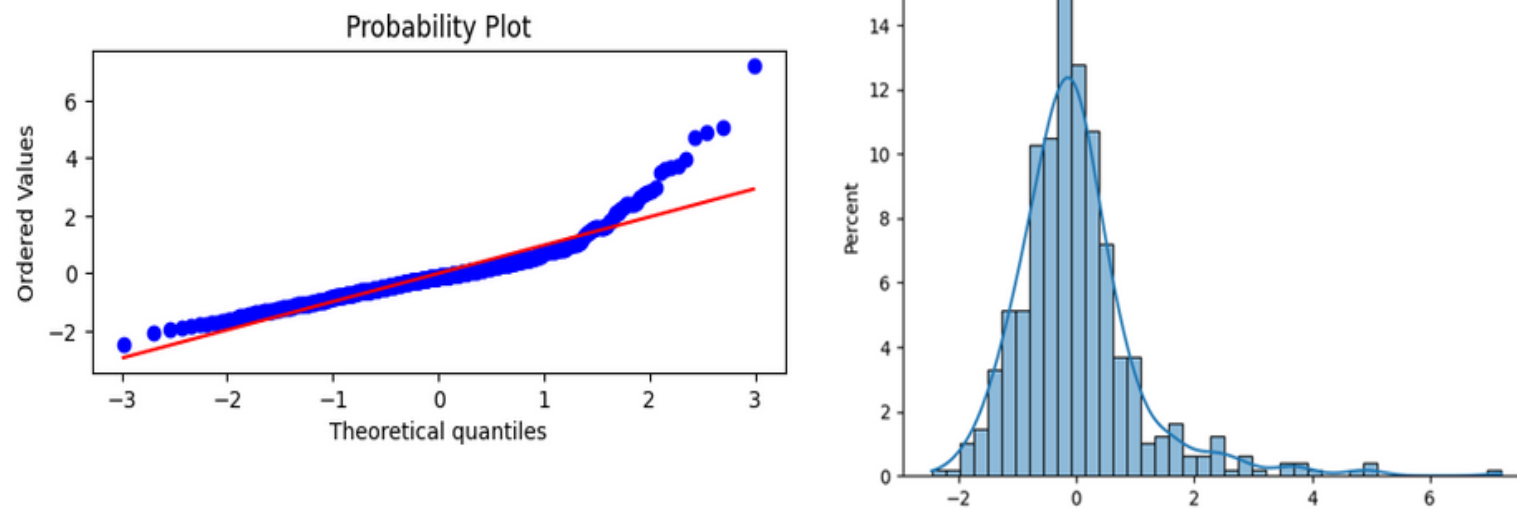


Рисунок В.9 – Графики нормального распределения остатков для модели №1

0.9556362252097914

Рисунок В.10 – Результаты теста Дарбина-Уотсона для модели №1

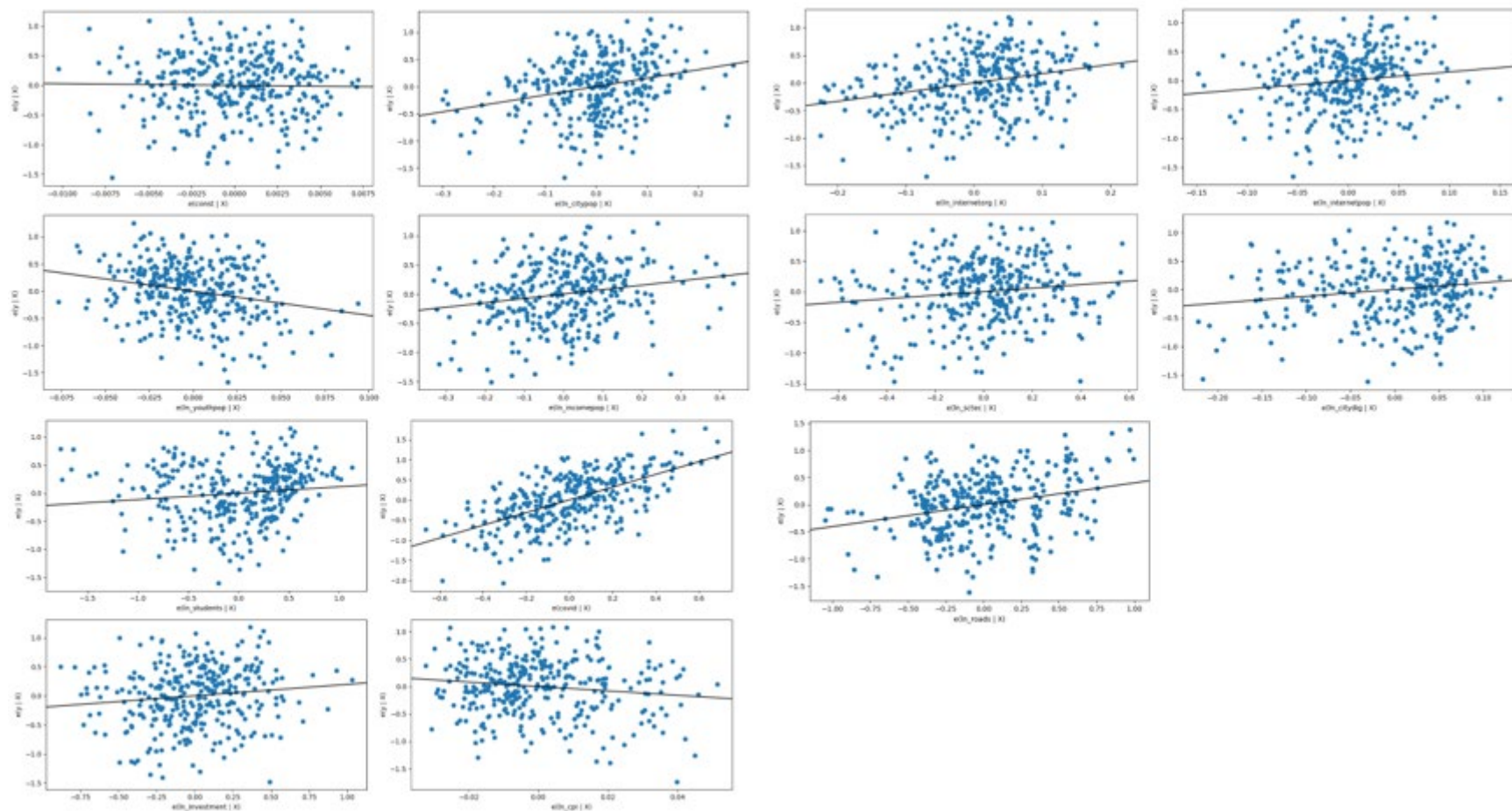


Рисунок В.11 – Графики частичной регрессии для модели №6

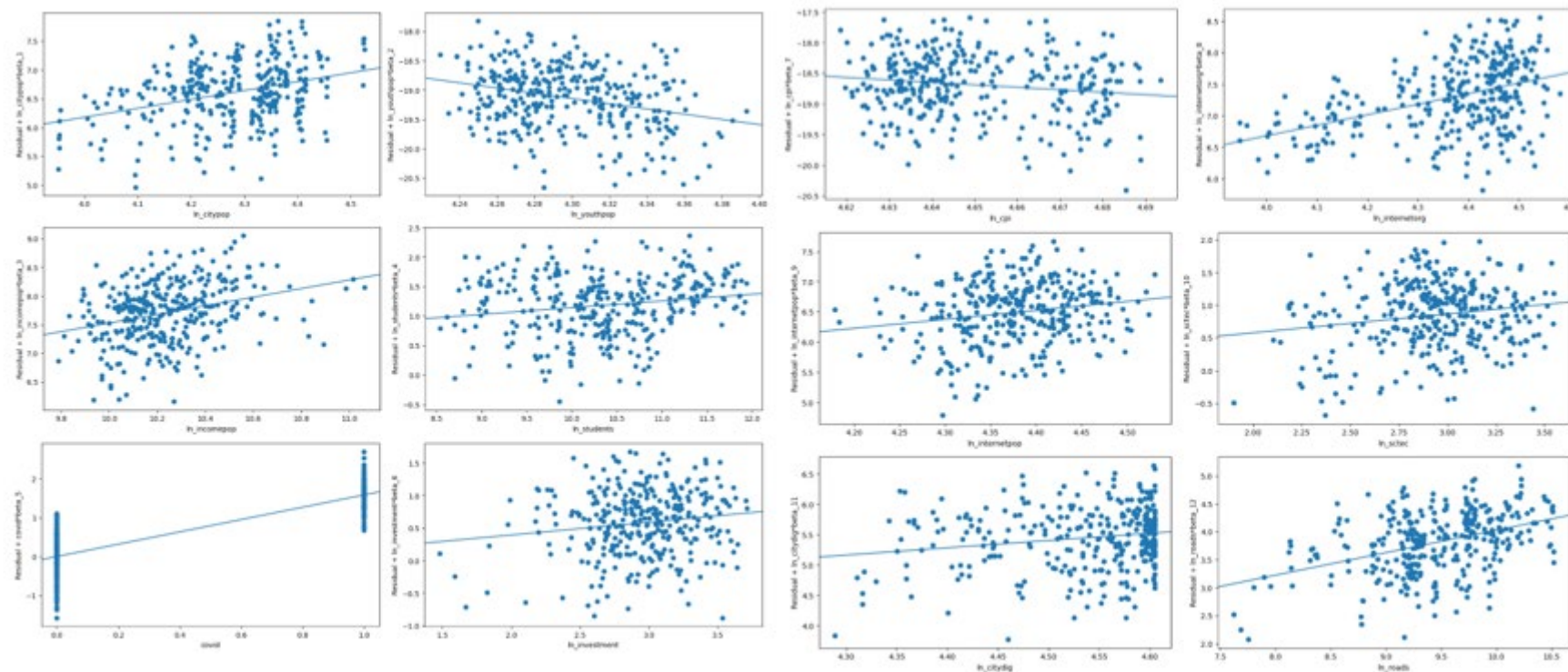


Рисунок В.12 – Графики частичных остатков для модели №6

RandomEffects Estimation Summary			
Dep. Variable:	ln_eshare	R-squared:	0.7873
Estimator:	RandomEffects	R-squared (Between):	0.5743
No. Observations:	330	R-squared (Within):	0.8031
Date:	Wed, May 24 2023	R-squared (Overall):	0.7302
Time:	21:07:10	Log-likelihood	-157.54
Cov. Estimator:	Unadjusted		
		F-statistic:	97.786
Entities:	55	P-value	0.0000
Avg Obs:	6.0000	Distribution:	F(12,317)
Min Obs:	6.0000		
Max Obs:	6.0000	F-statistic (robust):	97.786
		P-value	0.0000
Time periods:	6	Distribution:	F(12,317)
Avg Obs:	55.000		
Min Obs:	55.000		
Max Obs:	55.000		

Parameter Estimates						
	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
const	6.8479	10.428	0.6567	0.5119	-13.669	27.365
ln_citypop	1.6904	0.5501	3.0730	0.0023	0.6081	2.7726
ln_youthpop	-0.0836	1.7567	-4.6015	0.0000	-11.540	-4.6273
ln_incompop	0.8763	0.3296	2.6590	0.0082	0.2279	1.5248
ln_students	0.0940	0.0991	0.9406	0.3435	-0.1010	0.2890
covid	1.5042	0.1155	13.028	0.0000	1.2770	1.7313
ln_investment	0.1656	0.1007	1.6445	0.1011	-0.0325	0.3637
ln_cpi	-3.4902	1.3826	-2.5243	0.0121	-6.2104	-0.7699
ln_internetorg	1.5233	0.3248	4.6893	0.0000	0.8842	2.1624
ln_internetpop	1.8241	0.6474	2.8176	0.0051	0.5503	3.0979
ln_sctec	0.1317	0.1953	0.6744	0.5006	-0.2525	0.5159
ln_citydig	1.5022	0.5835	2.5745	0.0105	0.3542	2.6503
ln_roads	0.4144	0.1442	2.8750	0.0043	0.1308	0.6980

PanelOLS Estimation Summary			
Dep. Variable:	ln_eshare	R-squared:	0.8333
Estimator:	PanelOLS	R-squared (Between):	-9.5718
No. Observations:	330	R-squared (Within):	0.8333
Date:	Wed, May 24 2023	R-squared (Overall):	-2.4831
Time:	21:07:28	Log-likelihood	-105.48
Cov. Estimator:	Unadjusted		
		F-statistic:	109.59
Entities:	55	P-value	0.0000
Avg Obs:	6.0000	Distribution:	F(12,263)
Min Obs:	6.0000		
Max Obs:	6.0000	F-statistic (robust):	109.59
		P-value	0.0000
Time periods:	6	Distribution:	F(12,263)
Avg Obs:	55.000		
Min Obs:	55.000		
Max Obs:	55.000		

Parameter Estimates						
	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
const	101.44	41.421	2.4490	0.0150	19.880	183.00
ln_citypop	13.731	4.9305	2.7848	0.0057	4.0223	23.439
ln_youthpop	-37.552	5.8401	-6.4300	0.0000	-49.051	-26.053
ln_incompop	-1.2249	0.8893	-1.3774	0.1696	-2.9759	0.5261
ln_students	-0.4566	0.6079	-0.7512	0.4532	-1.6536	0.7403
covid	1.6714	0.1689	9.8983	0.0000	1.3389	2.0039
ln_investment	0.0776	0.1086	0.7143	0.4757	-0.1363	0.2914
ln_cpi	1.5487	1.6038	0.9656	0.3351	-1.6092	4.7066
ln_internetorg	0.9216	0.3739	2.4649	0.0143	0.1854	1.6578
ln_internetpop	0.3972	0.7321	0.5426	0.5879	-1.0444	1.8388
ln_sctec	-0.3902	0.3309	-1.1793	0.2393	-1.0418	0.2613
ln_citydig	0.8635	0.8579	1.0065	0.3151	-0.8258	2.5527
ln_roads	0.1860	0.9838	0.1891	0.8502	-1.7510	2.1231

Рисунок В.13 – Модели МНК со случайными и фиксированными эффектами