

Федеральное государственное автономное образовательное учреждение высшего  
образования «Национальный исследовательский университет «Высшая школа  
экономики»

Санкт-Петербургская Школа экономики и менеджмента  
Образовательная Программа "Экономика"

## КУРСОВАЯ РАБОТА

На тему

"Применение гравитационной модели к анализу миграций в Российской империи"

Выполнил студент группы БЭК-182  
Соснин Юрий Алексеевич

Научный руководитель:  
ст.преп. Куга Яков Тойвович

Санкт-Петербург  
2021 г.

# Содержание

1	Введение	3
2	Анализ внутренней миграции	4
2.1	Гравитационная модель . . . . .	4
2.2	Миграция в развивающихся странах . . . . .	6
2.3	Внутренние миграции в 19 веке . . . . .	8
3	Экономика Российской империи	12
4	Данные, модель и результаты	15
4.1	Данные . . . . .	15
4.2	Описательные статистики . . . . .	16
4.3	Гипотезы и модель . . . . .	18
4.4	Результаты . . . . .	20
5	Заключение	26
A	Таблицы	30
B	Онлайн-приложение	31
C	Карты	32

# 1 Введение

Связь экономического развития и миграции – большой и сложный вопрос современной экономики развития. Тем не менее, можно с уверенностью сказать, что миграции и экономическое развитие тесно связаны между собой.

Фундаментальные причины миграции – это разница в полезностях индивидов, принимающих решение о переезде, но на макро-уровне ее можно объяснить с помощью характеристик регионов. Некоторые свойства территорий, такие как, например, высокая безработица, выступают «выталкивающими», push-факторами для региона-источника миграции; низкая безработица в регионе-реципиенте, с другой стороны, это pull-фактор, «притягивающий» потенциальных мигрантов. Аналогия с гравитацией неслучайна – наверное, самая популярная модель, применяемая к миграциям населения – это гравитационная модель, прямая аналогия с законом Ньютона. Оценка гравитационной модели на данных о двухсторонних взаимодействиях позволяет определить причины, стоящие за решениями о миграции и выборе ее направления.

Моя работа посвящена внутренней миграции в Российской империи в конце 19 века. Последние исследования собирают все больше информации об экономике России перед Первой мировой войной (Cheremukhin и др., 2017; Markevich & Zhuravskaya, 2018) и о региональном распределении экономического развития (Lindert & Nafziger, 2014; Markevich, 2019). И хотя исследованию причин внутренней миграции в других великих державах 19 века посвящено некоторое направление литературы о исторической демографии, работы о России очень немногочисленны (Anderson, 1980; Leasure & Lewis, 1968), и им недостает экономической методологии и современного взгляда.

Я использую данные Всероссийской переписи населения 1897 года (Н. А. Тройницкий, 1897) о миграции и некоторых других показателях, а также набор других исторических источников, чтобы выяснить, какие региональные факторы оказывали влияние на внутреннюю миграцию в конце 19 века в России. Среди региональных факторов, наиболее часто считающихся детерминантами миграций и индустриальных экономиках – демографические (плотность населения, уровень урбанизации), экономические (характеристики рынка труда и индустриальный выпуск) и социальные (уровень грамотности). Я проверяю гипотезы относительно этих переменных для миграции в города и уезды империи.

Работа структурирована следующим образом. Сначала я рассказываю о гравитационной модели и способах ее оценки в секциях 1.1 и 1.2. Далее – о ее применении к внутренней миграции в современных развивающихся странах и к историческим миграциям в европейских странах в 19 веке. После этого, я кратко рассказываю о Российской империи в 19 веке и задаю исследовательский вопрос. В практической части сначала идет обсуждение данных и их описательные статистики, затем – я обсуждаю гипотезы и модель, и наконец, результаты. Заканчивается работа заключением, обсуждающим ограничения результатов и дальнейшее обсуждение.

## 2 Анализ внутренней миграции

### 2.1 Гравитационная модель

Впервые основные гипотезы пространственного моделирования миграций предложил в 1885 году британский географ и статистик Эрнст Георг Равенштайн в знаменитой статье *The Laws of Migration* в *Journal of the Royal Statistical Society*. (Ravenstein, 1885) Автор проанализировал данные британских переписей 1871 и 1881 годов, определяя основные направления и силу потоков внутренней миграции между Англией, Шотландией и Ирландией, а также между графствами внутри этих королевств. Равенштайн не предложил математической модели миграции, его анализ напоминает скорее современную описательную статистику, однако на основе своих наблюдений он сформулировал несколько гипотез – «законы миграции», легшие в основу будущих моделей.

Во-первых, автор отмечает роль расстояния в миграционных потоках: большинство мигрантов преодолевают небольшое расстояние. Кроме того, «... we must take into account the number of natives of each county that furnishes the migrants, as also the population of the towns or districts which absorb them» – сила миграционного потока зависит от населения, готового принимать в нем участие, от способности регионов отдавать и принимать мигрантов.

Во-вторых, миграции имеют определенную пространственную структуру. В первую очередь, города привлекают жителей близлежащей сельской местности, потом – жителей соседних графств, и, наконец, жителей остальной страны (морское сообщение, при этом, также способствует миграции). С другой стороны, даже между самыми дальними регионами есть двустороннее миграционное «сообщение». Каждый поток имеет противоположный.

В-третьих, Равенштайн отмечает, что основная причина миграции – рынок труда. Жители Британии чаще переселялись в торговые и промышленные центры, с более широкими возможностями для трудоустройства.

Наконец, мигранты имеют особые характеристики: молодые, не имеющие детей, и жители деревень более предрасположены к миграции, чем, соответственно, более взрослые, жители города и семьи с детьми. Кроме того, в случае Британии 19 века, женщины более мобильны, чем мужчины.

Работа Равенштайна оказала значительное влияние на будущие работы демографов, социологов, и, в конечном счете, экономистов, определив основные гипотезы для моделей миграции и пространственного взаимодействия в целом. Первый закон миграции в понимании Равенштайна, например, излагает положения базовой гравитационной модели.

В дальнейшем работа над моделями миграции населения продолжилась. Stouffer расширил определение расстояния до «intervening opportunities» – всех препятствующих переезду факторов (транспортные расходы, информация о месте

назначения, законодательные барьеры, культурные отличия и так далее) (Stouffer, 1940). Позднее была предложена гравитационная модель – модель двухстороннего пространственного социального взаимодействия - основанная на Ньютоновском законе гравитации (Stewart, 1948; Zipf, 1946). Изначально включавшая в себя только населения источника и назначения и расстояние, она была расширена включением push- и pull-факторов – доходов, климата, качества социальной политики, безработицы, и т.д. (Lee, 1966). Одновременно с этим развивалась применение модели к международной торговле, где она оказалась также успешной (De Benedictis & Taglioni, 2011).

Современная форма базовой гравитационной модели миграции выглядит так:

$$M_{ij} = \beta_0 P_i^{\beta_1} P_j^{\beta_2} D_{ij}^{\beta_3} + \varepsilon_{ij} \quad (1)$$

Здесь,  $M$  – число переселенцев из региона  $i$  в регион  $j$ ;  $P_i$  и  $P_j$  – население региона-источника и региона-назначения,  $D_{ij}$  – расстояние,  $\varepsilon_{ij}$  – случайный фактор. Расширенная модель кроме населения будет также включать всевозможные притягивающие и отталкивающие характеристики регионов, а расстояние не обязательно должно быть географическим.

Сегодня, гравитационная модель миграции – уже не просто результат эмпирических наблюдений за миграционными потоками. В отличие от школы «социальной физики», к которой принадлежали первые авторы модели, экономистам необходима связь между эмпирическими макро-законами и решениями индивидов, максимизирующих полезность, микро-основания. Основное микроэкономическое объяснение гравитационных моделей с помощью решений индивидов – модель максимизации случайной полезности (Random utility maximization, RUM) (см., например, Veine и др., 2016). Основанием модели RUM лежит полезность от переезда в другой регион, определенная как разность детерминированной полезности нахождения в том регионе и издержек на переезд, плюс случайный ненаблюдаемый фактор. Предположение о распределении этого фактора определяет ожидаемую вероятность того, что выбор определенного альтернативного региона окажется оптимальным выбором индивида. От этого можно перейти к ожидаемой суммарной миграции между регионами, которая оказывается положительно зависящей от привлекательности региона-реципиента и его способности отдавать мигрантов, и отрицательно – от привлекательности региона-источника и его «доступности», что соответствует выражению гравитационной модели.

Есть несколько методов оценки гравитационной модели, самым очевидным из которых является МНК с логарифмическим преобразованием.

$$\ln(M_{ij}) = \beta_0 + \beta_1 \ln(P_i) + \beta_2 \ln(P_j) + \beta_3 \ln(D_{ij}) + \varepsilon_{ij} \quad (2)$$

Несмотря на удобство и простоту интерпретации, такая спецификация имеет ряд недостатков (Silva & Tenreyro, 2006). Во-первых, логарифм от нуля не определен, тогда как нулевые миграционные потоки часто составляют заметную часть наблюдений. Два основных способа решения этой проблемы – это исключение нулевых наблюдений и

прибавление к  $M$  небольшой константы, как правило 1 или 0,5. При этом исключение наблюдений с нулями смещает оценку, но и размер константы невозможно определить теоретически, что снижает точность и обоснованность оценок логарифмической модели.

Если и другая, теоретическая проблема с логарифмированием моделей с постоянной эластичностью, или мультипликативных моделей, отмеченная Silva и Tenreyro. Чтобы случайный фактор в логарифмической спецификации не зависел от регрессоров, нужно сделать насчет него очень строгие предположения, что в большинстве случаев нереалистично – поэтому Silva и Tenreyro не рекомендуют логарифмировать гравитационные модели.

Вместо этого, авторы предложили использовать метод Poisson pseudo-maximum-likelihood. Этот метод естественным образом справляется с нулями в данных, оценивая модель в ее мультипликативной форме с помощью метода максимального правдоподобия. Предполагается, что миграционные потоки распределены по Пуассону с условным средним  $\mu_{ij}$ , которое является функцией объясняющих переменных. Формально это определяется следующим образом:

$$Pr[M_{ij}] = \frac{\exp(-\mu_{ij})\mu_{ij}^{M_{ij}}}{M_{ij}!}, \quad M_{ij} = (0, 1, \dots)$$

$$\mu_{ij} = \exp(\beta_0 + P_i\beta_1^T + P_j\beta_2^T + D_{ij}\beta_3^T)$$

Такая оценка стала в последнее время бенчмарком в исследованиях с применением гравитационной модели, все больше авторов предпочитают ее логарифмам.

В целом, спецификации гравитационной модели выходит далеко за рамки расширенной гравитационной модели. Использование панельных данных позволяет применить фиксированные эффекты для источника, реципиента и для пар регионов; random intercept позволяет при этом оценить и неизменные факторы регионов. Тем не менее, моя работа – в силу доступности данных – кросс-секционная, и применять я буду PPML estimator для расширенной модели.

## 2.2 Миграция в развивающихся странах

Итак, гравитационные модели активно и успешно применяются к анализу внутренней миграции в современных странах (Andrienko & Guriev, 2004; Etzo, 2008; Guriev & Vakulenko, 2015; Molloy и др., 2011). Основной фокус обзора литературы, однако, будет направлен на литературу о миграциях в 19 веке – но прежде чем перейти к таким работам (которые, по большей части, написаны в 80-х годах прошлого века), стоит все-же обратиться к современным применениям гравитационных моделей миграции. Будут рассмотрены примеры статей, посвященных современным развивающимся странам. Предположение о том, что современных развивающиеся страны похожи на страны Европы в период индустриализации, конечно, далеко от того, чтобы

быть истинным в полной мере, однако многие процессы можно назвать схожими – в частности, урбанизация, развитие промышленности и концентрация населения – происходили в Европе 18-19 веков, происходят и в развивающихся странах в недавнем прошлом. Я не буду в этом обзоре обращать внимание на все, или даже на самые популярные; основная цель – составить представление о том, какие гипотезы проверяют авторы и какие переменные в целом используют.

В большинстве работ о внутренней миграции с применением гравитационной модели авторы стремятся, оценив расширенную гравитационную модель, определить региональные детерминанты миграции. В работах обсуждается, какие характеристики регионов в целом оказывают влияние на перемещение населения, а также проверяются гипотезы относительно конкретных переменных. В качестве данных используют панели, основанные на переписях населения; оценивают гравитационные модели либо обычным МНК, либо с помощью PPML – эта техника становится все популярнее в последние годы.

Несколько примеров работ, посвященных миграции в развивающихся странах. Так, Flores и др. проверяют влияние на миграции либерализации внешней торговли в результате вступления Мексики в НАФТА (Flores и др., 2013). Amara и Jemmali выясняет, как на внутренние миграции в Тунисе влияют безработица, количество доступных вакансий и расходы домохозяйств (Amara & Jemmali, 2018). Dang и др. обнаруживают, что мигранты стремятся в более индустриализованные и богатые регионы, а государственные программы, призванные этому помешать почти не оказывают никакого влияния (Dang и др., 1997); van Lottum и Marks приходят к таким же выводам для Индонезии (van Lottum & Marks, 2012). Giang и др. находят положительную связь между внутренними миграциями и качеством государственного управления во Вьетнаме (Giang и др., 2020). Kone и др. определяют, что границы штатов в Индии негативно влияют на миграцию в масштабе районов – правда, авторы контролируют только на национальные различия, но не на экономические факторы (Kone и др., 2018).

Wajdi и др. и Royuela и Ordóñez используют гравитационные модели для подтверждения гипотезы о концентрации населения. И в Эквадоре, и в Индонезии, развивающихся странах, миграции направлены в крупные города, то есть регионы с высоким уровнем урбанизации и высоким населением. Население, таким образом, концентрируется в нескольких агломерациях, города и регионы с уже высоким населением продолжают расти, притягивая жителей менее населенных районов сильнее, чем это происходит в обратном направлении. В обоих случаях, тем не менее, этот тренд ослабевает с течением времени.

Etzo (2008) выделяет 4 основные группы переменных расширенной гравитационной модели:

1. Гравитационные переменные
2. Экономические переменные

### 3. Переменные рынка труда

### 4. Переменные «окружающей среды»

Рассмотренные статьи подтверждают такие приоритеты. Экономические и социально-экономические характеристики регионов являются главными, помимо населения и расстояния, переменными в моделях для развивающихся стран. Во многих работах, так или иначе, в качестве предикторов выступают плотность населения, уровень урбанизации, структура занятости или уровень безработицы. Немногие, правда, применяют географические контроли – отчасти потому, что рассмотренные страны, как правило, являются довольно однородными в этом плане, чего нельзя будет сказать о России.

Во всех рассмотренных работах, миграции оказываются направлены из менее развитых регионов в более развитые – мигрантов «притягивают» возможности трудоустройства в более урбанизированных и промышленно развитых областях, что полностью соотносится с микроэкономическими обоснованиями гравитационных моделей.

## 2.3 Внутренние миграции в 19 веке

Интересно заметить, что статья Равенштайна 1885, положившая начало пространственным количественным исследованиям внутренней миграции, сегодня может считаться и исторической: объект работы – Британия 19 века, миграции исследуются в контексте эпохи индустриальной экономики до Первой мировой войны, источник – данные переписей.

Позднее, исследователи снова обратили внимание на европейскую историю: существует направление исторической демографии, занимающееся анализом внутренней миграции в странах Европы для этого времени. Тем не менее, сейчас эти работы по большей части устаревши. Кроме того, несмотря на доступность исторических данных, таких работ немного: McQuillan в *Economic Factors and Internal Migration: The Case of Nineteenth-Century England* (1980), описывая состояние литературы по проблеме внутренней миграции в историческом контексте, отмечает, что ей уделено мало внимания. Это было справедливо в 80-х годах прошлого века, отчасти справедливо и по сей день. Все работы – это либо демографические работы, описывающие основные направления и особенности миграционных потоков, либо исторические case-study, описывающие особенности миграции небольших групп; немного внимания уделено экономическим исследованиям причин межрегиональной миграций в масштабе стран.

Сам McQuillan не применяет тогда уже разработанную гравитационную модель – для определения детерминант миграции он рассматривает корреляции между темпом out-migration и объясняющими переменными, такими, как доля занятых в сельском хозяйстве, доля нищих, доход и географическая площадь. Данные, используемые в



исследовании – переписи населения в Англии 1851-1901 годов. Методология таких работ, конечно, вызывает вопросы. Современные методы (в т.ч. гравитационные модели) позволяют ответить на вопросы о макро-причинах миграции с большей достоверностью.

Greenwood и Thomas (1973) применяют гравитационную модель с сельскохозяйственными доходами к миграциям в Англии и Уэльсе в 19 веке. Авторы приходят к выводу, что внутренняя миграция отрицательно зависела от сельскохозяйственных доходов в регионе-источнике, положительно – от возможностей индустриального трудоустройства в регионе-направления, и стандартно негативно – от расстояния. В противоположность гипотезе Равенштайна, миграция на большое расстояние не оправдывалась индустриальным развитием: в целом, города привлекали трудовые ресурсы из близлежащих регионов с низкими сельскохозяйственными доходами.

Другая статья о миграции в Великобритании, *Internal migration in England, 1818-1839*, фокусируясь на другом периоде, периоде более ранней индустриализации, отвечает на похожие вопросы (Nicholas & Shergold, 1987). Используя данные о сосланных в Австралию жителей Англии (которые оказываются относительно репрезентативны относительно всего населения Англии), исследователи заключают, что различия в степени урбанизации, как прокси для возможностей трудоустройства, негативно влияли на поток миграции между регионами; уровень грамотности, как в регионе-источнике, так и в регионе-реципиенте, имел положительный знак; сельскохозяйственные доходы показывали такие же зависимости, как и для более поздних периодов.

Использование необычных микро-данных для гравитационной модели нередко помогает исследователям в более ранние периоды истории. Charpentier и Gallic, например, используют данные генеалогических деревьев с сайта Geneanet, чтобы отследить миграции во Франции в начале 19 века. В целом, исследования миграции в период до индустриализации или ранней индустриализации в Европе составляют значительную часть литературы. Кроме Англии и Франции, McCants анализирует внутреннюю миграцию во Фрисландии, историческом регионе на севере Нидерландов в период 1750-1805 годов (McCants, 1992). Автор, правда, не использует расширенную гравитационную модель, зато доказывает, что представление о резком увеличении темпов внутренней миграцией неверно: население Европы было довольно мобильно и в более ранние периоды, особенно на короткие расстояния.

Как можно заметить, основное внимание в литературе, в продолжение традиции Равенштайна, уделено Великобритании. При этом стоит отметить также работы, посвященные Германии, Испании и Италии в конце 19 – начале 20 века (Hochstadt, 1981; Hogan & Kertzer, 1985; Silvestre, 2005).

Исследования внутренней миграции в Европе 19 века не сильно отличаются от работ о современных развивающихся странах. Проверяются те же гипотезы: мигранты предпочитают перемещаться в более богатые, промышленно развитые и

урбанизированные регионы, что в 19 веке, что сейчас.

Что касается России, есть всего несколько работ, посвященных количественному исследованию внутренних миграций в Российской империи конца 19 – начала 20 века. Статья Leasure и Lewis и монография Барбары Андерсон (Anderson, 1980) наиболее близки к теме данной работы: они анализируют межрегиональные миграции по данным переписи 1897 года с помощью макро-моделей и пытаются выявить характеристики регионов, влиявшие на миграционные потоки.

Leisure и Lewis в 1968 году предприняли попытку, используя данный переписи, определить, были ли мигранты в Российской империи более склонны переезжать в более развитые регионы с большими доходами. (Leasure & Lewis, 1968) Оказалось, что степень урбанизации, уровень грамотности и процент занятых в промышленности связан с более высокими уровнями миграции в регион, а занятость в сельском хозяйстве – с более низким уровнем.

Исследовательский вопрос и результаты схожи с вопросами и результатами в литературе, касающейся других стран, однако методы по современным меркам оставляют желать лучшего. В частности, авторы не используют гравитационную модель, наиболее подходящую для такого рода исследования, но, более того, даже не рассматривают все имеющиеся данные. «Because it was much too laborious to calculate and analyze in-migration to each guberniia from each of the other eighty-eight guberniias or foreign countries ..., it was decided to consider for each guberniia only the top five suppliers of migrants». Кроме того, как метод определения зависимости авторы использовали ранговую корреляцию Пирсона, вместо регрессий, а также не включили достаточного количества переменных контроля. Несмотря на недостатки, эта статья – по-прежнему одна из немногих, посвященных вопросу миграции в Российской империи.

Барбара Андерсон в монографии 1980 года выдвигает нетипичную гипотезу: выбор направления миграции зависит не столько от экономических показателей, сколько от «схожести» места назначения с изначальным местом жительства мигранта. Даже неграмотный потенциальный мигрант из региона с высоким уровнем грамотности будет иметь более «современные» взгляды на жизнь и будет предпочитать переезд в регион с высоким уровнем грамотности; то же самое и с естественным приростом населения и промышленным выпуском. Автор тоже не применяет гравитационную модель, вместо этого оценивая регрессии для миграции в отдельные регионы из других регионов, с расстоянием, грамотностью, естественным приростом населения, выпуском промышленности и производительностью земли в качестве зависимых переменных. Б. Андерсон делит страну на несколько больших регионов в зависимости от их «modernity», и анализирует миграции внутри них. В целом, гипотезы подтверждаются: более современные регионы привлекали мигрантов из более современных регионов в смысле уровня грамотности, структуры трудоустройства и прироста населения (он рассматривается как мера стадии демографического перехода и сам по себе оказывается притягивающим фактором, что противоречит результатам для других стран). Тем

не менее, малое количество использованных данных (по 20-50 наблюдений в одной регрессии), недостаточное количество переменных контроля (культурные различия между регионами или географические характеристики регионов не используются) ставят под сомнение результаты исследования.

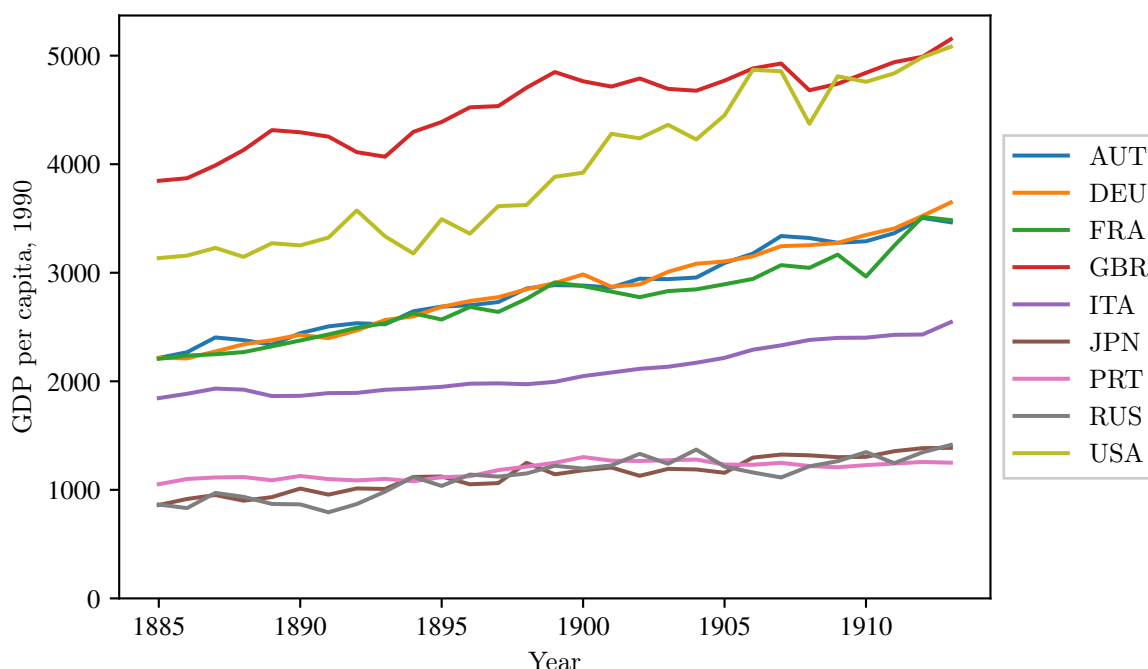
И Leasure и Lewis, и Anderson отмечают наличие других данных о миграции в поздней Российской империи – это данные о переселении крестьян в Сибирь и на Дальний восток незадолго до и во время реформ Столыпина (Турчанинов, 1910, 1916). Они не в полной мере подходили для двух предыдущих работ, не будут использоваться и в данной курсовой, однако на них основана самая заметная работа о внутренней миграции в РИ – статья А. Маркевича и др. Property rights and internal migration.

Chernina и др. исследуют связь внутренней миграции и ликвидности земли. Реформа Столыпина, кроме прочего, позволила крестьянам некоторых центральных губерний выходить из передельных общин, получая право собственности на свою землю, и, соответственно, возможность ее продать. Рассматривая реформу как естественный эксперимент, повлиявший на провинции с передельными общинами и не повлиявший на другие губернии и области, авторы, с помощью панельных данных о переселении крестьян за Урал в 1901-1914, показывают, что изменение прав собственности на землю увеличило миграцию независимо от других аспектов реформы.

### 3 Экономика Российской империи

Российская империя в конце 19 века – интенсивно развивающаяся страна с динамически развивающейся экономикой, только что вступившая в стадию того, что называют современным экономическим ростом. В 1880-1900 годах ВВП, по расчетам Gregory (1983), рос на 2% в год – нормальные значения для того времени, темпы, сопоставимые с развитыми экономиками того времени. Тем не менее, как в абсолютных значениях ВВП, так и в плане выпуска промышленности, экономика Российской империи была довольно отсталой – на уровне Японии или Португалии, с общим ВВП на душу населения значительно ниже Германии, Великобритании, или США.

Рис. 1: ВВП на душу населения, 1990\$ (Inklaar и др., 2018)



Интересно, кроме того, региональное распределение валового продукта. Имело место значительное неравенство между регионами империи. Санкт-Петербург и Москва имели подушевой ВВП, вполне сопоставимый с крупнейшими экономиками мира того времени, тогда как беднейшие регионы Центральной Азии отличались от столицы по этому показателю в 12 раз (Markevich, 2019).

Вообще, отдельно стоит сказать об источнике таких данных – статье А. Маркевича *A Regional Perspective on the Economic Development of the late Russian Empire* (2019). Автор выясняет, какие факторы влияли на региональную экономическую развития в Российской империи в конце 19 века. Для этого он, по данным 1897 года, реконструирует региональный валовый продукт, факторную производительность в промышленности и сельском хозяйстве, государственные расходы на образование, оборону, и так далее, для всех регионов империи, и проверяет гипотезы о влиянии на них географических и

Таблица 1: Подушевой продукт, 1897, 1990\$ (Markevich, 2019)

Российская империя	1122
Санкт-Петербургская губерния	6826
Тургайская область	567
Великобритания	4428
США	3780
Португалия	1182
Япония	1062

институциональных переменных.

Оказывается, что и география, и институты являлись причинами различий в экономическом развитии регионов (что соответствует результатам для других стран). Выход к морю, количество свободной земли и наличие минеральных ресурсов могут объяснить положительную вариацию, в то время как наследие крепостного права и другие институциональные характеристики приводили к более низким экономическим результатам регионов.

Эта статья – не только ответ на вопрос о том, чем же объяснялось региональное неравенство или вклад в литературу об экономическом росте, но и огромная работа по реконструкции экономических показателей России в конце 19 века. Вместе с демографическими и социальными данными Всероссийской переписи населения 1897 года, доступными в оцифрованном виде на сайте проекта Электронного архива Российской исторической статистики ([ristat.org](http://ristat.org), Маркевич и Кесслер, 2020), эти данные позволяют взглянуть на пространственное распределение многих показателей.

Как отмечает Markevich, Российская империя не имела строгого промышленного развитого центра и неразвитой периферии, дело обстоит скорее наоборот. Крупные промышленные центры располагались неравномерно и чаще по окраинам – Лодзь, Варшава, Баку, Донбасс, и, наконец, Москва и Санкт-Петербург – тому доказательства. Кроме того, наиболее благоприятные сельскохозяйственные земли – это юг Украины и юг Западной Сибири (Томская губерния), не самые центральные регионы. В то же время, регионы центральной России, наиболее сильно зависевшие от труда крепостных, были перенаселены, неэффективны и в целом бедны.

На основании этих наблюдений, я задаю вопрос: какие факторы могут объяснить внутреннюю миграцию? Оказывалось ли перенаселение региона фактором, вытесняющим мигрантов в другие места? Ехали ли мигранты преимущественно в промышленные городские центры, или наиболее значимым было переселение крестьян в плодородные окраины? В конце концов, были ли детерминанты внутренней миграции сопоставимы в России и Великобритании или Германии, того же периода? На эти вопросы я постараюсь ответить в данной работе.

Несмотря на то, что Markevich (2019) – это препринт, и данные о ВРП и продуктивности труда еще нигде автором не опубликованы, и многие другие необходимые переменные не удалось найти, эта работа – моя попытка рассказать с

помощью гравитационной модели о внутренних миграциях в Российской империи в конце 19 века.

## 4 Данные, модель и результаты

### 4.1 Данные

Вообще, качество имперской статистики считается хорошим, количество данных достаточным для исследования самых разных вопросов экономической истории (Markevich, 2019).

Основной источник статистических данных о поздней Российской империи – Первая всеобщая перепись населения Российской империи 1897 года. Переписные тетради для каждого региона (кроме Финляндии) содержат множество данных, среди которых – численность населения, распределения населения по родному языку, вероисповеданию, грамотности, возрасту, семейному состоянию. Есть также таблицы распределения «неместных уроженцев по месту рождения»: число человек, живущих в данном регионе, родившихся в другом регионе (или иностранном государстве) – по регионам рождения, городам и уездам в регионе нахождения.

Данные о распределении жителей регионов по месту рождения позволяют составить матрицу двухстороннего пространственного взаимодействия, из которой получить зависимую переменную гравитационной модели – число человек, приехавших из каждого из регионов в каждый другой регион.

Конечно, у этого источника есть существенные недостатки, накладывающие ограничения на потенциальные результаты исследования.

Во-первых, это кросс-секционные данные, перепись 1897 года – первая и единственная всеобщая перепись населения в Российской империи. Существуют панельные данные о переселении крестьян в Сибирь и на дальний восток в 1901-1914, во времена Столыпинской реформы (см. Chernina и др., 2014), однако, для всей империи таких данных нет, тем более что нет и ежегодных данных о грамотности, урбанизации, и т.д.

Во-вторых, в отличие от современных данных о миграциях, использующихся в современных гравитационных моделях, данные о количестве переселившихся носят «пожизненный» характер: известно только о месте рождения и месте пребывания, неизвестно, когда человек совершил миграцию. Современные данные, напротив, как правило представляют миграцию в какой-то определенный период – год, промежуток времени с предыдущей переписи, и т.д. Эта особенность, на самом деле, вызывает опасения насчет эндогенности. Высокая средняя грамотность на Сахалине – это pull-фактор, причина повышенной миграции, или ее следствие – просто более грамотные люди в среднем мигрируют чаще (а заселение Сахалина как раз происходило в конце 19 века)? Из-за природы данных, выводы о причинно-следственных связях и интерпретацию результатов нужно делать с осторожностью.

Наконец, как уже было упомянуто, данных о характеристиках регионов – используемых и как переменные контроля, и как переменные интереса для проверки гипотез – доступно гораздо меньше, чем того хотелось бы. Ценные результаты оценок

ВРП А. Маркевича еще, к сожалению, не опубликованы.

Данные о миграции из переписи я собираю сам; большинство других показателей переписи доступны на [ristat.org](http://ristat.org) (Маркевич & Кесслер, 2020). Я также использую данные многих других источников, обработанные для своих целей другими исследователями – долю крепостных (оригинальный источник – Статистический временник Российской империи, 1866), длину железных дорог и продуктивность сельского хозяйства из Markevich и Zhuravskaya, 2018; структуру занятости, цены на землю и доход крестьян – из Lindert и Nafziger, 2014 (оригинальный источник – Опыт приблизительного исчисления народного дохода по различным его источникам и по размерам в России, 1906). Подробнее об источниках см. таблицу 6.

## 4.2 Описательные статистики

В 1897 году 9% населения проживало вне региона своего рождения. 5,7% сельского населения страны были мигрантами, в то время как доля неместных в городах приближалась к 30% (из города или уезда они уехали – неизвестно). В среднем один мигрант преодолевал расстояние в 860 километров; для городских мигрантов это расстояние было меньше, 718 километров, а для сельских – 975.

Важнейшее направление миграции в абсолютных значениях – Санкт-Петербург, Москва, Кубанская область (примерно сегодняшний Краснодарский край) и Томская губерния, включавшая в себя плодородные земли Алтая (см. карту 3). При этом, если отдельно учесть мигрантов в городах и уездах, это сразу обращает внимание на 2 основных направления миграции – в крупные города (Москву и Санкт-Петербург) и на плодородные земли юга (Кубань и Алтай), о которых, в частности, говорится и в литературе (Anderson, 1980; Leasure & Lewis, 1968). Миграция на Дальний восток (восточнее Томской губернии), в абсолютных значениях, пока незначительна – до реформы Столыпина еще 9 лет.

В Санкт-Петербургской губернии в 1897 году 42% жителей родились в других регионах империи. Из 2,1 миллионов ее жителей, 800 тысяч были мигрантами в первом поколении. В Москве эта доля составляла 27%.

На Дальнем востоке, на Сахалине процент приезжих равен почти 70, в Амурской области – 47, в Приморской области – 39. В Кубанской области 31% жителей приехали из других регионов, в Черноморской губернии – 57%, в Томской – 32%.

Что касается отдачи населения, абсолютными рекордсменами являются регионы центральной России и северо-восточной Украины. От 15 до 20% населения (население – на 1897 год, а уезжали люди и раньше, так что это строго говоря не out-migration rate, но, тем не менее, показательно) покинули Калужскую, Тульскую, Рязанскую, Ярославскую, Тверскую, Курскую, Воронежскую, Полтавскую губернии (см. карту 2). Это хорошо соответствует выводам А. Маркевича, касающихся относительной отсталости центральной России в поздней империи: перенаселение, наследие крепостного права, превалирование общинной собственности на землю – все



это делало эти регионы довольно бедными и, соответственно, отталкивало население (Markevich, 2019).

Значительна также внутренняя миграция в Польше. Основные направления миграции в Привислинских губерниях – это Варшава, и Петроковская губерния, в которой находился на тот момент крупный центр текстильной промышленности город Лодзь. Стоит отметить и внутреннюю миграцию евреев внутри зоны оседлости – например, в Лодзи, Одессе и Варшаве евреи составляли до трети населения.

Несколько слов нужно сказать о некоторых из объясняющих переменных: населении, плотности населения, уровне урбанизации, выпуске промышленности (см. приложение В).

Всего в Российской империи за исключением Финляндии в 1897 году проживало 129,2 млн. человек. Среднее значение населения по регионам – 1,4 миллиона человек. Самыми населенными в абсолютных значениях регионами империи были губернии Украины, Черноземья, а также Самарская, Саратовская, Вятская и Пермская губернии. Что касается плотности населения, наибольшей она была в Привислинских губерниях, а также в Киевской, Подольской и Московской губерниях. Очень низкой, соответственно, она была в Сибири, на Севере, а также на Кавказе.

Общий уровень урбанизации по всей Российской империи составлял 15%. Среди регионов выделяются столичная Санкт-Петербургская губерния (67%), Московская (46%), Варшавская (43%), Херсонская (28%) и Петроковская (36%) губернии – Санкт-Петербург, Москва, Варшава, Одесса и Лодзь в таком порядке – крупнейшие города Российской империи в 1897. Стоит отметить, что относительно урбанизированным оказывается Дальний Восток: 23 и 27 процентов в Приморской и Амурской областях.

Средний процент грамотных среди всего населения империи составлял 29,3%, по регионам среднее значение составляет 18%. Самая высокая доля грамотных в прибалтийских губерниях – от 60 до 70%, самая низкая – в Закавказье и Средней Азии – около 3%.

Самыми промышленно развитыми регионами по выпуску на душу населения были Петроковская (текстиль), Санкт-Петербургская, Бакинская (нефть), Московская, Владимирская губернии, губернии Прибалтики, а также Екатеринославская губерния (региона Донбасса) и Амурская область. Стоит отметить, что в том числе и по подсчетам Маркевича, регионы Дальнего востока не были исключительно аграрными регионами фронта – они были довольно богатыми и развитыми. Если рассматривать долю населения, занятого в промышленности, лидеры остаются примерно такими же, однако перестают выделяться Донбасс и Баку – судя по всему, их высокий выпуск в рублях был заслугой небольшого числа рабочих, относительно всего населения.

Таблица 2: Гипотезы

Переменная	Гипотеза
population <sub>i</sub> , population <sub>j</sub> distance	population <sub>i</sub> < population <sub>j</sub> $\approx -1$
urbanization	+
density	–
literacy	+
industry	+

### 4.3 Гипотезы и модель

Согласно исследовательскому вопросу о факторах, влиявших на внутреннюю миграцию, были выдвинуты следующие гипотезы.

Первая группа гипотез касается населения и переменных, связанных с населением. Как отмечается в работах о развивающихся странах, на ранних стадиях экономического развития крупные регионы еще сильнее «укрупняются».

В контексте гравитационной модели (1) это означает взаимную зависимость коэффициентов при населении регионов.

$$M_{ij} - M_{ji} = M_{ij} \left[ 1 - \frac{M_{ji}}{M_{ij}} \right] = M_{ij} \left[ 1 - \left( \frac{P_i}{P_j} \right)^{\beta_2 - \beta_1} \right] \quad (3)$$

Если  $\beta_2 > \beta_1$ , то  $M_{ij} > M_{ji}$  (Poot и др., 2016), для каждой пары регионов более населенные регионы притягивают мигрантов относительно сильнее, и вся система стремится к концентрации. Таким образом, если Российская империя находилась на ранних этапах индустриализации и испытывала сильную rural-to-urban миграцию, коэффициент при населении региона-реципиента будет выше, чем при населении региона-источника.

С одной стороны, плотность населения выше в городских центрах, тогда высокие ее значения будут скоррелированы с привлекательностью для мигрантов; с другой стороны, высокая плотность населения может в среднем быть выталкивающим фактором: если в центральных регионах наблюдалось перенаселение, а мигранты стремились из них уехать, это должно отразиться в значимом положительном коэффициенте плотности населения региона-источника.

Во-вторых, так как основная движущая сила миграций – рынок труда, переменные, связанные с лучшими возможностями трудоустройства должны быть pull-факторами. Для современных стран с современными данными информация о зарплатах и безработице и их различия между регионами – это основные объясняющие переменные гравитационных моделей, помимо обязательных расстояния и населения. Для исторических миграций такие данные как правило недоступны, поэтому приходится использовать другие переменные. Для стран в период ранней индустриализации мера возможностей городского трудоустройства – это социально-экономическое

развитие. Грамотность, урбанизация и выпуск промышленности должны притягивать мигрантов. Это подтверждается историческими исследованиями для Британии, Испании, Германии. (Hochstadt, 1981; Nicholas & Shergold, 1987; Silvestre, 2005)

Итоговая модель выглядит так:

$$M_{ij} = \beta_0 S_i D_j X_{ij} + \epsilon_{ij} \quad (4)$$

$S$  – факторы предложения, характеристики региона-источника,  $D$  – факторы спроса, характеристики региона-назначения,  $X$  – характеристики пар регионов.

$$\begin{aligned} S_i &= population_i^{\beta_1} \times urbanization_i^{\beta_2} \times density_i^{\beta_3} \times literacy_i^{\beta_4} \times industry_i^{\beta_5} \\ D_j &= population_j^{\beta_6} \times urbanization_j^{\beta_7} \times density_j^{\beta_8} \times literacy_j^{\beta_9} \times industry_j^{\beta_{10}} \\ X_{ij} &= distance_{ij}^{\gamma_1} \times e^{\gamma_2 border_{ij}} \times e^{\gamma_3 sea_{ij}} \times \Delta climate_{ij}^{\gamma_4} \times \Delta language_{ij}^{\gamma_5} \end{aligned}$$

Я оцениваю эту модель в мультипликативной форме с помощью PPML. (Silva и Tenreiro, 2006, см. ??). В такой формулировке модель нужно записать так:

$$Pr[M_{ij}] = \frac{\exp(-\mu_{ij}) \mu_{ij}^{M_{ij}}}{M_{ij}!}, \quad M_{ij} = (0, 1, \dots) \quad (5)$$

Где  $M$  – миграция между регионом  $i$  и регионом  $j$ ,  $\mu$  – среднее распределения Пуассона, зависит от переменных следующим образом,  $S$ ,  $D$  и  $X$  - уже в линейном виде:

$$\mu_{ij} = \exp(\beta_0 + S_i \beta_1^T + D_j \beta_2^T + X_{ij} \beta_3^T) \quad (6)$$

Кроме переменных интереса, необходимых для проверки гипотез, в модели присутствуют некоторые географические и пространственные переменные контроля. (см. таблицу 6) Различия в характеристиках климата – среднегодовой температуре, влажности воздуха, количестве морозных и дождливых дней, среднем давлении насыщенного пара – должны отражать, в первую очередь, относительную пригодность региона для сельского хозяйства, а также нежелание мигрантов переезжать в незнакомый и непривычный климат.

В модель входят также абсолютные различия в долях населения, говорящих на восточнославянских языках (русский, украинский и белорусских), на польском, немецком и «еврейских» языках для отражения широкой национальной неоднородности Российской империи. Миграции между регионами с разными культурами, при прочих равных, будет ниже.

Наконец, кроме непосредственно географического расстояния между административными центрами регионов, в модель добавлены дамми на общую границу и наличие выхода к морю (кроме Северного Ледовитого океана), для лучшего

моделирования физических и экономических издержек мигрантов на переезд.

Многие переменные недоступны для всей страны, есть их расчеты только для Европейской ее части – многие исследования поздней Российской империи опираются только на относительно однородные европейские губернии. Для последовательности результатов, я оцениваю также полную модель только для внутриевропейских миграций, а также модели с добавлением других переменных: доли крепостных в 1868 году (*serfs*), длины железных дорог в 1897 (*railway*), продуктивности выращивания пшеницы (*grain<sub>prod</sub>*), цены земли за 1900-1902 (*land<sub>price</sub>*) и средних доходов крестьянских хозяйств (*peasant<sub>inc</sub>*) (источники данных см. в таблице 6).

Большинство работ так или иначе включают сельскохозяйственные доходы как объясняющую переменную, поэтому я предполагаю, что они будут значимыми детерминантами миграции: большие доходы (как и большая продуктивность сельского хозяйства) должны притягивать мигрантов. Длина железных дорог может отражать как общее экономическое развитие (развитость инфраструктуры), так и относительно меньшие издержки миграции. Более дорогая земля может как притягивать, так и снижать миграции в соответствующие регионы.

Эти результаты, тем не менее, не являются основными для моей работы: все-таки исключение Сибири и Дальнего востока серьезно меняет картину внутренних миграций и не позволит оценить факторы, на нее влиявшие.

#### 4.4 Результаты

Основные результаты можно увидеть в таблице 3. Модель (1) – самая базовая гравитационная модель, только с расстоянием и населением. (2) и (3) – это полные модели со всеми переменными, отличающиеся спецификацией промышленного выпуска; (4) – иллюстративная модель без плотности населения.

Коэффициент при переменной расстояния значим во всех спецификациях и равен примерно  $-1$  в полных моделях (3) и (4) и  $-1,2$  в базовой модели. Этот результат полностью сходится с литературой: коэффициент, как правило, в базовой модели выше, так как забирает на себя часть дисперсии пропущенных переменных. Кроме того, дамми на общую границу между регионами (*borders*) оказывается предсказуемо значимым.

Оценка влияния населения дает смешанные результаты. Оба коэффициента значимы и находятся в своих естественных промежутках, однако интересно не это, а их взаимная зависимость. Предполагалось, что население региона-реципиента окажется более важным, по сравнению с населением региона-источника, и это так в полных моделях. Однако интересно заметить, что такой результат не получается, если «пропустить» плотность населения (*density*). Сама плотность населения оказывается очень важным объясняющим фактором, значимым для обеих сторон взаимодействия: высокая плотность в регионе-источнике выталкивает лишних жителей, одновременно снижая и миграции в регионы с высокой плотностью. Намного ниже коэффициент для плотности населения по сравнению с сельскими мигрантами у мигрантов в города

Таблица 3: Результаты

	(1)	(2)	(3)	(4)
	Simple	Industry shares	Industry outputs	No density
population_i	0.927***	0.783***	0.761***	0.914***
population_j	0.523***	1.314***	1.364***	0.657***
distance	-1.204***	-1.040***	-1.028***	-0.808***
borders		1.008***	1.008***	1.190***
sea_i		-0.0356	-0.0690	-0.395***
sea_j		0.0853	0.0678	0.652***
literacy_i		0.233**	0.248**	0.371***
literacy_j		0.105	0.165*	-0.177**
urbanization_i		-0.0990	-0.160	-0.0625
urbanization_j		1.045***	1.274***	0.708***
density_i		0.442***	0.426***	
density_j		-0.879***	-0.857***	
industry_share_i		-0.107		0.00337
industry_share_j		0.176		0.0417
serfs_i		-0.0132	-0.0168	-0.0228
serfs_j		-0.166***	-0.163***	-0.194***
eastslavic_abs_diff		-0.0597***	-0.0659***	-0.0740***
polish_abs_diff		-0.0187	-0.0240	-0.0259
jewish_abs_diff		-0.0474	-0.0516	-0.0774**
german_abs_diff		0.0594**	0.0568**	0.0594**
capital_i		-0.212	-0.145	-0.138
capital_j		0.232	0.164	0.292
tmp_diff		-0.00757	-0.0238	-0.224***
wet_diff		-0.118***	-0.110***	-0.130***
pre_diff		0.0292***	0.0286***	0.00886*
frs_diff		-0.0913	-0.110	-0.226**
vap_diff		-0.0361	-0.0549	-0.0994
industry_pc_i			-0.0170	
industry_pc_j			-0.0696	
Constant	-5.141***	-11.90***	-11.82***	-8.316***
Observations	7832	7656	7656	7656
$R^2$	0.192	0.503	0.497	0.453

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ 

(таблица 4), и это логично – плотность населения эквивалентна количеству свободной земли, это важно для сельского населения и неважно для городского. Те же результаты

сохраняются, если вместо общей плотности населения взять только сельскую, а также если добавить произведение плотности и доли городских жителей (результаты не приведены).

Таким образом, сравнивая результаты базовой регрессии с результатами расширенной гравитационной модели, можно заключить, что в среднем более населенные регионы были источником миграции, а менее населенные – местом назначения, однако этот эффект пропадает, если включить в модель другие характеристики, позволяющее лучше моделировать решения индивидов. Грубо говоря, мигранты уезжали из более населенных регионов, но по другим причинам. Это соответствует некоторым свидетельствам. Российская империя не была на той стадии развития, на которой население концентрируется в крупных центральных регионах; население крупных центральных европейских губерний активно уезжало из них во второй половине 19 века, освобожденное от крепостного права. Еще согласно известной гипотезе Ключевского и модели Домара, крепостное право само по себе нужно, чтобы удерживать крестьян от переезда на окраины большой страны, обеспечивая доходы землевладельческой элиты (Domar, 1970). Судя по всему, это оставалось справедливо вплоть до его отмены в 1861 году, и последствия этого отражает моя модель. Несмотря на то, что переменная, отражающая наследие крепостного права – *serfs* – незначима для региона-источника, для региона-реципиента она значимо отрицательна.

Уровень урбанизации принимающего региона оказывается значим для всех моделей, в том числе и для миграции в уезды. Это косвенно свидетельствует о том, что урбанизация действительно неплохо отражает «развитость» губернии – сельские мигранты (которые не обязательно крестьяне, мануфактурные производства на селе также существовали в это время) таким образом могли получать доступ к более развитым ранкам и инфраструктуре.

Уровень грамотности оказывается незначим как характеристика региона-реципиента. Это не соответствует некоторым из работ, включавших уровень грамотности как объясняющую переменную (Nicholas & Shergold, 1987). Положительное влияние грамотности в регионе-источнике можно объяснить тем, что грамотные люди с более хорошими представлениями о мире более склонны к миграции. С другой стороны, грамотность это также и мера экономического развития, тогда получается, что жители более богатых регионов были более склонны к миграции.

В противоречии с гипотезами Ревенштайна, результатами для других европейских стран 19 века, и результатами Leasure и Lewis промышленность, судя по результатам, никак не влияла на внутреннюю миграцию в Российской империи. Коэффициент оказывается значим только в регрессии с сельскими мигрантами, однако меняет знак, если заменить долю занятых в промышленности на выпуск на душу. Для российских мигрантов гораздо важнее были демографические и климатические условия принимающего региона (урбанизация и плотность населения), чем присутствие еще не столь развитой промышленности.

Таблица 4: Urban-Rural breakdown

	(1)	(2)	(3)
	mig_total	mig_rural	mig_urban
population_i	0.783***	0.841***	0.801***
population_j	1.314***	1.339***	1.182***
distance	-1.040***	-0.940***	-1.170***
borders	1.008***	1.342***	0.654***
sea_i	-0.0356	0.251	-0.260*
sea_j	0.0853	0.0167	0.0932
literacy_i	0.233**	0.131	0.267**
literacy_j	0.105	0.171	0.00894
urbanization_i	-0.0990	-0.0344	-0.141
urbanization_j	1.045***	0.368**	1.929***
density_i	0.442***	0.708***	0.177*
density_j	-0.879***	-1.087***	-0.595***
industry_share_i	-0.107	-0.255	0.0969
industry_share_j	0.176	0.432***	-0.207
serfs_i	-0.0132	0.00709	-0.0613
serfs_j	-0.166***	-0.180***	-0.0729***
eastslavic_abs_diff	-0.0597***	-0.0669**	-0.0555**
polish_abs_diff	-0.0187	-0.0562	0.0183
jewish_abs_diff	-0.0474	-0.0173	-0.0700**
german_abs_diff	0.0594**	0.0710**	0.0393*
capital_i	-0.212	-0.506*	0.00675
capital_j	0.232	-0.0632	-0.0587
tmp_diff	-0.00757	0.0229	-0.0158
wet_diff	-0.118***	-0.143***	-0.118***
pre_diff	0.0292***	0.0407***	0.0140**
frs_diff	-0.0913	-0.0590	-0.0256
vap_diff	-0.0361	0.0518	0.0189
Constant	-11.90***	-15.79***	-9.009***
Observations	7656	7656	7656
$R^2$	0.503	0.323	0.576

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Переменные контроля ведут себя предсказуемо, показывая ожидаемые знаки. Выход к морю не оказывается значим ни для одной из сторон миграционного взаимодействия, как и сам факт того, расположен ли в регионе один из крупнейших городов страны. Различия в национальном составе областей снижали между ними миграцию. Различия

по климатическим переменным склоняло миграцию в пользу лучшего климата.

Таким образом, гипотезы частично не подтвердились. Более высокая доля городского населения действительно значимо привлекала мигрантов, в том числе и сельских. Промышленный выпуск, при этом, никак не был связан с миграцией, как и уровень грамотности, что противоречит результатам для других стран. Население, несмотря на значительный рост крупнейших городов, не концентрировалось в нескольких регионах, а наоборот, разъезжалось из регионов с высокой плотностью населения (преимущественно, центральных).

Перейдем к результатам дополнительных моделей только для Европейской части страны, с добавлением переменных, рассчитанных только для этих регионов. Как видно из таблицы 5, оценка изначальной модели только для европейских губерний дает примерно те же результаты, с одним исключением – выпуске промышленности все же оказывается значимым pull-фактором. Добавление новых переменных, однако, не приносит ожидаемого эффекта. Предполагалось, что продуктивность земли должна притягивать как минимум сельских мигрантов, а различия в доходах крестьян и цене земли выступать как значительные драйверы миграции. На деле, многие из переменных оказываются незначимыми, а некоторые и вовсе имеют значимые «неправильные» знаки.

Стоит отметить относительно низкий  $R^2$  всех получившихся моделей. Одно из известных свойств гравитационных моделей – очень высокий «fit», высокий процент объясненной дисперсии. Wajdi и др., например, получают  $R^2$ , равный 0,9; в то же время, значения в результатах моих оценок разнятся от 0,36 до 0,61. Возможно, и это нужно отметить снова, это следствие недостаточного количества необходимых переменных (тест Рамсея также не проходит для 2 и 3 степени). Тем не менее, как я уже упоминал, получить данные о других переменных – увы – очень сложно или пока невозможно, так что это возможность для будущего продолжения работы над этой темой.



Таблица 5: Европейская Россия

	(1)	(2)	(3)	(4)
	All regions	Only European	Added vars, urban	Added vars, rural
population_i	0.783***	0.543***	0.534	0.940***
population_j	1.314***	1.300***	1.333***	0.918***
distance	-1.040***	-0.999***	-1.172***	-0.942***
borders	1.008***	1.169***	0.678**	1.751***
sea_i	-0.0356	-0.246	-0.00427	0.471*
sea_j	0.0853	0.462**	0.251	0.661**
literacy_i	0.233**	0.209	0.224	0.297
literacy_j	0.105	0.0635	0.112	0.149
urbanization_i	-0.0990	-0.153	-0.141	0.0517
urbanization_j	1.045***	0.741***	1.599***	-0.426*
density_i	0.442***	0.320***	0.533**	0.295
density_j	-0.879***	-0.657***	-0.415*	-1.111***
industry_share_i	-0.107	0.0635	0.123	-0.247
industry_share_j	0.176	0.417**	0.111	0.657***
serfs_i	-0.0132	-0.0433	-0.0857	0.0844
serfs_j	-0.166***	-0.0419	-0.0640	0.00672
capital_i	-0.212	-0.123	-0.328	-0.875**
capital_j	0.232	0.211	-0.00117	0.349
railway_i			0.00529	-0.0723
railway_j			-0.0766	0.136*
grain_prod_i			-0.105	-0.118
grain_prod_j			-0.0818	-0.269**
land_price_i			0.0489	0.404*
land_price_j			-0.244	0.344*
peasant_inc_i			-0.267	0.665
peasant_inc_j			-2.251**	-0.113
Constant	-11.90***	-8.987**	8.515	-16.13
Observations	7656	2450	2256	2256
$R^2$	0.503	0.615	0.631	0.606

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

## 5 Заключение

Я собрал данные, оценил модель и получил результаты.

Результаты соотносятся с литературой.

Интерпретация результатов, конечно, имеет ограничения, связанные с природой данных. Как уже было упомянуто, данные доступны только в кросс-секционном формате, а в переписи учитываются «пожизненные» миграции: это не позволяет применить более робастные панельные методы оценки гравитационной модели, а также не позволяет оценить динамику реакции мигрантов на меняющиеся экономические условия, и, соответственно, сделать выводы о причинно-следственной связи миграции и объясняющих переменных.

Кроме того, добыча данных оказалась очень сложной задачей. Многие показатели требуют извлечения из документов – я ограничился только доступными в электронном виде данными. Многие ключевые и интересные показатели (как ВРП или уровень безработицы) или не существуют, или не существуют в открытом доступе. Тем не менее, я считаю, что и в получившемся виде модель имеет ценность как базовая модель миграции в России в конце 19 века.

В дальнейшем, особенно с использованием данных Андрея Маркевича (Markevich, 2019), с помощью моей модели можно будет оценивать влияние других переменных: например, институциональных ограничений, государственных расходов и программы и т.д.

## Список литературы

- Amara, M. & Jemmali, H. (2018). Deciphering the Relationship Between Internal Migration and Regional Disparities in Tunisia. *Social Indicators Research*, 135(1), 313–331.
- Anderson, B. A. (1980). *Internal Migration During Modernization in Late Nineteenth-Century Russia*. Princeton University Press.
- Andrienko, Y. & Guriev, S. (2004). Determinants of interregional mobility in Russia. *Economics of Transition and Institutional Change*, 12(1), 1–27.
- Beine, M., Bertoli, S. & Moraga, J. F.-H. (2016). A Practitioners' Guide to Gravity Models of International Migration. *The World Economy*, 39(4), 496–512.
- Charpentier, A. & Gallic, E. (2018). Internal Migrations in France in the Nineteenth Century (arXiv № 1807.08991).
- Cheremukhin, A., Golosov, M., Guriev, S. & Tsyvinski, A. (2017). The Industrialization and Economic Development of Russia through the Lens of a Neoclassical Growth Model. *The Review of Economic Studies*, 84(2), 613–649.
- Chernina, E., Castañeda Dower, P. & Markevich, A. (2014). Property rights, land liquidity, and internal migration. *Journal of Development Economics*, 110, 191–215.
- Dang, A., Goldstein, S. & McNally, J. (1997). Internal Migration and Development in Vietnam. *International Migration Review*, 31(2), 312–337.
- De Benedictis, L. & Taglioni, D. (2011). The Gravity Model in International Trade. B L. De Benedictis & L. Salvatici (Ред.), *The Trade Impact of European Union Preferential Policies: An Analysis Through Gravity Models* (с. 55–89). Springer.
- Domar, E. D. (1970). The Causes of Slavery or Serfdom: A Hypothesis. *The Journal of Economic History*, 30(1), 18–32.
- Etzo, I. (2008). Internal migration: a review of the literature (MPRA Paper). University Library of Munich, Germany.
- Flores, M., Zey, M. & Hoque, N. (2013). Economic Liberalization and Contemporary Determinants of Mexico's Internal Migration: An Application of Spatial Gravity Models. *Spatial Economic Analysis*, 8(2), 195–214.
- GeoHack. (б.г.). Получено 1 марта 2021 с [geohack.toolforge.org](https://geohack.toolforge.org)
- Giang, L. T., Nguyen, C. V. & Nguyen, H. Q. (2020). The impacts of economic growth and governance on migration: evidence from Vietnam. *The European Journal of Development Research*, 1–35.
- Global Price and Income History Group. (б.г.). Получено 1 мая 2021 с [gpih.ucdavis.edu](https://gpih.ucdavis.edu)
- Greenwood, M. J. & Thomas, L. B. (1973). Geographic labor mobility in nineteenth century England and Wales. *The Annals of Regional Science*, 7(2), 90–105.
- Gregory, P. R. (1983). *Russian National Income, 1885–1913*. Cambridge University Press.
- Guriev, S. & Vakulenko, E. (2015). Breaking out of poverty traps: Internal migration and interregional convergence in Russia. *Journal of Comparative Economics*, 43(3), 633–649.

- Harris, I., Osborn, T. J., Jones, P. & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, 7(1), 109.
- Hochstadt, S. (1981). Migration and Industrialization in Germany, 1815-1977. *Social Science History*, 5(4), 445–468.
- Hogan, D. P. & Kertzer, D. I. (1985). Migration Patterns During Italian Urbanization, 1865-1921. *Demography*, 22(3), 309–325.
- Inklaar, R., de Jong, H., Bolt, J. & van Zanden, J. (2018). Rebasing 'Maddison': new income comparisons and the shape of long-run economic development (GGDC Research Memorandum GD-174). Groningen Growth & Development Centre, University of Groningen.
- Kone, Z. L., Liu, M. Y., Mattoo, A., Ozden, C. & Sharma, S. (2018). Internal borders and migration in India. *Journal of Economic Geography*, 18(4), 729–759.
- Leasure, J. W. & Lewis, R. A. (1968). Internal Migration in Russia in the Late Nineteenth Century. *Slavic Review*, 27(3), 375–394.
- Lee, E. S. (1966). A theory of migration. *Demography*, 3(1), 47–57.
- Lindert, P. H. & Nafziger, S. (2014). Russian Inequality on the Eve of Revolution. *The Journal of Economic History*, 74(3), 767–798.
- Markevich, A. (2019). A Regional Perspective on the Economic Development of the late Russian Empire (SSRN Scholarly Paper № ID 2555273). Social Science Research Network. Rochester, NY.
- Markevich, A. & Zhuravskaya, E. (2018). The Economic Effects of the Abolition of Serfdom: Evidence from the Russian Empire. *American Economic Review*, 108(4-5), 1074–1117.
- McCants, A. (1992). Internal Migration in Friesland, 1750-1805. *The Journal of Interdisciplinary History*, 22(3), 387–409.
- McQuillan, K. (1980). Economic Factors and Internal Migration: The Case of Nineteenth-Century England. *Social Science History*, 4(4), 479–499.
- Molloy, R., Smith, C. L. & Wozniak, A. (2011). Internal Migration in the United States. *Journal of Economic Perspectives*, 25(3), 173–196.
- Nicholas, S. & Shergold, P. R. (1987). Internal migration in England, 1818–1839. *Journal of Historical Geography*, 13(2), 155–168.
- Poot, J., Alimi, O., Cameron, M. P. & Maré, D. C. (2016). The Gravity Model of Migration: The Successful Comeback of an Ageing Superstar in Regional Science (SSRN Scholarly Paper № ID 2864830). Social Science Research Network. Rochester, NY.
- Ravenstein, E. G. (1885). The Laws of Migration. *Journal of the Statistical Society of London*, 48(2), 167–235.
- Royuela, V. & Ordóñez, J. (2018). Internal migration in a developing country: A panel data analysis of Ecuador (1982-2010). *Papers in Regional Science*, 97(2), 345–367.

- Sablin, I., Kuchinskiy, A., Korobeinikov, A., Mikhaylov, S., Kudinov, O., Kitaeva, Y., Aleksandrov, P., Zimina, M. & Zhidkov, G. (2018). Transcultural Empire: Geographic Information System of the 1897 and 1926 General Censuses in the Russian Empire and Soviet Union.
- Silva, J. M. C. S. & Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics*, 88(4), 641–658.
- Silvestre, J. (2005). Internal migrations in Spain, 1877–1930. *European Review of Economic History*, 9(2), 233–265.
- Stewart, J. Q. (1948). Demographic Gravitation: Evidence and Applications. *Sociometry*, 11(1/2), 31–58.
- Stouffer, S. A. (1940). Intervening Opportunities: A Theory Relating Mobility and Distance. *American Sociological Review*, 5(6), 845–867.
- van Lottum, J. & Marks, D. (2012). The determinants of internal migration in a developing country: quantitative evidence for Indonesia, 1930–2000. *Applied Economics*, 44(34), 4485–4494.
- Wajdi, N.,ADIOETOMO, S. M. & Mulder, C. H. (2017). Gravity Models of Interregional Migration in Indonesia. *Bulletin of Indonesian Economic Studies*, 53(3), 309–332.
- Zipf, G. K. (1946). The P1 P2\D Hypothesis: On the Intercity Movement of Persons. *American Sociological Review*, 11(6), 677–686.
- Маркевич, А. & Кесслер, Х. (2020). Электронный архив Российской исторической статистики, XVIII – XXI вв. <https://ristat.org/>
- Опыт приблизительного исчисления народного дохода по различным его источникам и по размерам в России. (1906). тип. П. П. Сойкина.
- Статистический временник Российской империи. (1866). Центральный статистический комитет Министерства внутренних дел.
- Тройницкий, А. Г. (1861). Крепостное население в России по 10-й народной переписи. Статистический отдел центрального статистического комитета.
- Тройницкий, Н. А. (1897). Первая всеобщая перепись населения Российской империи 1897 года. Центральный статистический комитет Министерства внутренних дел.
- Турчанинов, Н. В. (1910). Итоги переселенческого движения за время с 1896 по 1909 гг. (включительно). Переселенческое управление.
- Турчанинов, Н. В. (1916). Итоги переселенческого движения за время с 1910 по 1914 гг. (включительно). Переселенческое управление.

## А Таблицы

Таблица 6: Переменные

Переменная	Описание	Источник
mig_total	Миграции	Н. А. Тройницкий, 1897
distance	Расстояние между административными центрами	“GeoHack”, б.г.
population	Население	Маркевич и Кесслер, 2020
density	Плотность населения	
literacy	Уровень грамотности, доля населения	
urbanization	Доля городских жителей	
industry_pc	Промышленный выпуск на душу населения	
industry_share	Доля населения, занятого в промышленности	“Global Price and Income History Group”, б.г.
serfs	Доля крепостных в 1858	Markevich и Zhuravskaya, 2018; А. Г. Тройницкий, 1861
border	1, если пара регионов имеет общую границу	Sablin и др., 2018
sea	1, если есть выход к морю, кроме Северного Ледовитого океана	
capital	1, если Москва или Санкт-Петербург	
eastslavic	Доля говорящих на русском, украинском или белорусском	
polish	Доля говорящих на польском	
jewish	Доля говорящих на еврейских языках	
german	Доля говорящих на немецком	
tmp	Среднегодовая температура в 1901-1910, градусы Цельсия	Harris и др., 2020
wet	Среднее число дождливых дней в месяц в 1901-1910	
pre	Среднее количество осадков в 1901-1910, кг на кв.м.	
frs	Среднее число морозных дней в 1901-1910	
var	Среднее давление насыщенного пара в 1901-1910, мм рт.ст.	

Таблица 7: Описательные статистики

	Observations	Mean	Standard deviation	Min	Median	Max
population	89.0	1411685.63	825800.43	28113.00	1403901.00	3559220.00
density	89.0	32.92	27.77	0.08	30.99	130.00
urbanization	88.0	0.13	0.10	0.03	0.11	0.30
literacy	89.0	0.18	0.12	0.01	0.15	0.40
industry_pc	89.0	20.32	32.04	0.26	8.52	160.00
industry_share	89.0	0.09	0.06	0.01	0.08	0.30
east_slavic	89.0	55.69	38.51	0.63	70.82	99.00
polish	89.0	8.84	23.13	0.02	0.23	80.00
german	89.0	1.32	2.33	0.00	0.23	10.00
jewish	89.0	3.56	5.50	0.01	0.39	10.00
tmp	97.0	4.67	4.60	-13.82	4.90	10.00
wet	97.0	12.69	3.10	3.41	13.91	10.00
pre	97.0	43.63	12.93	7.69	44.97	90.00
frs	97.0	14.59	3.28	7.59	14.41	20.00
vap	97.0	7.37	1.60	2.75	7.68	10.00
serfs	89.0	0.24	0.28	0.00	0.06	0.30
railway	49.0	5.93	1.71	0.00	6.53	10.00
grain_prod	49.0	4.43	0.89	2.85	4.16	6.00
peasant_inc	50.0	470.38	64.94	393.54	471.21	820.00
land_price	49.0	81.01	46.49	15.00	73.00	180.00
mig_from	89.0	125815.20	107504.94	1138.00	120505.00	443740.00
mig_urban_from	89.0	56077.58	46905.27	783.00	49132.00	222500.00
mig_rural_from	89.0	69737.62	72558.82	355.00	52933.00	344050.00
mig_to	89.0	125815.20	153871.96	11794.00	70556.00	899260.00
mig_urban_to	89.0	56077.58	108560.97	1866.00	27968.00	827660.00
mig_rural_to	89.0	69737.62	91375.30	799.00	40845.00	574910.00

## В Онлайн-приложение

- [Карты для всех переменных](#)
- [Репозиторий с текстом работы, данными и кодом](#)

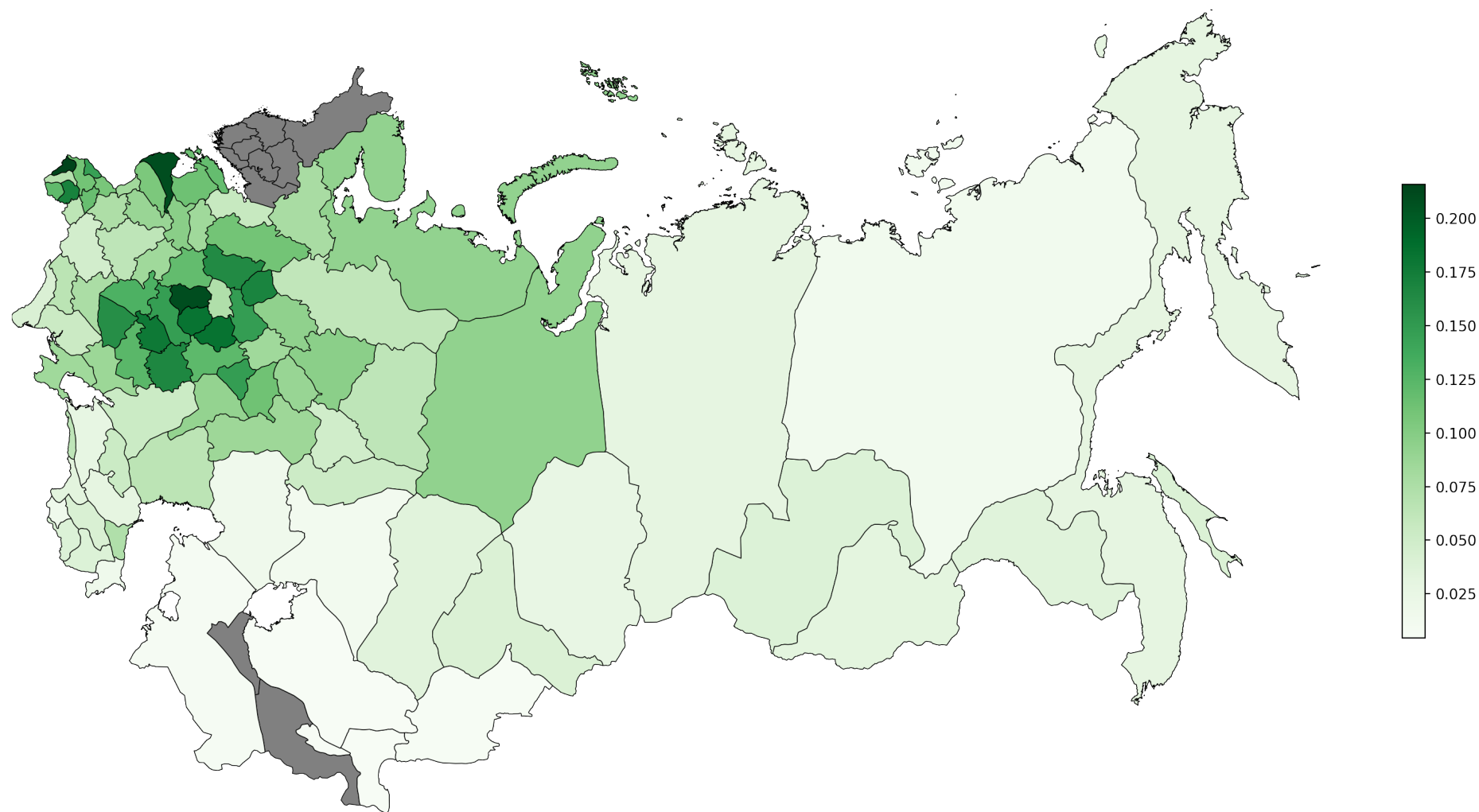


Рис. 2: Родившихся в регионе, но проживающих в других регионах, доля населения



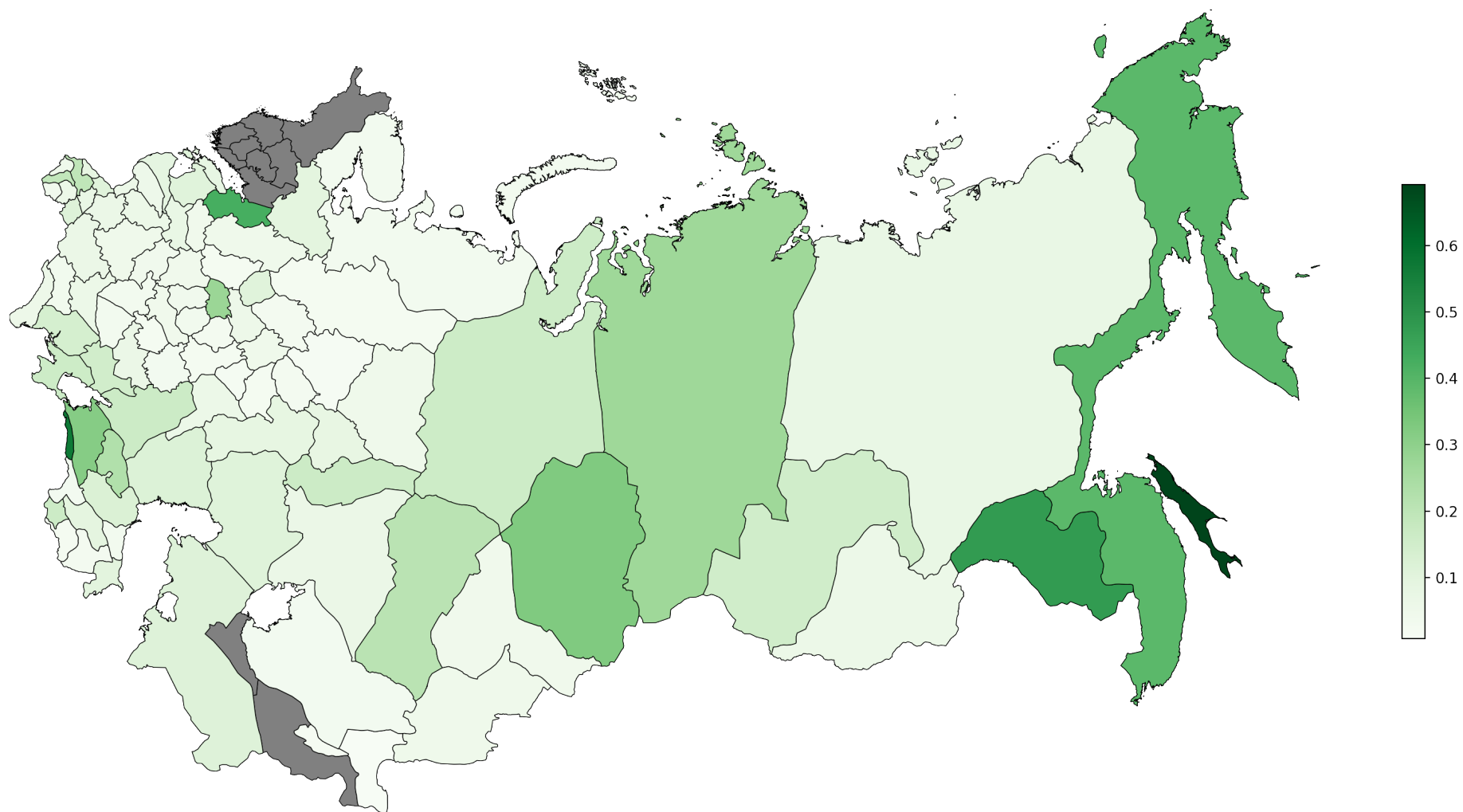


Рис. 3: Родившихся в других регионах, но проживающих в этом, доля населения