Семерикова Елена Вячеславовна. Использование пространственной зависимости в региональных исследованиях на примере анализа безработицы в России и Германии: диссертация ... кандидата Экономических наук: 08.00.13 / Семерикова Елена Вячеславовна;[Место защиты: ФГАОУВО Национальный исследовательский университет Высшая школа экономики], 2017.- 137 с.

**Содержание к диссертации**

Введение

**Глава 1 Пространственные эконометрические модели: особенности и применение 15**

1.1 Возникновение и развитие идей пространственной эконометрики 15

1.2 Взвешивающая матрица в моделях пространственной эконометрики 23

1.3 Обзор эмпирических исследований 27

1.3.1 Пространственный анализ региональных рынков труда 27

1.3.2 Использование пространственных эконометрических моделей для анализа российских региональных данных 31

**Глава 2 Применение пространственных эконометрических моделей для анализа региональной безработицы в России и Германии 36**

2.1 Выбор взвешивающей матрицы на примере региональной безработицы в России и Германии 36

2.1.1 Теоретические формулы для смещения пространственных эконометрических моделей 37

2.1.2 Данные и переменные 39

2.1.3 Результаты предварительного оценивания по реальным данным 40

2.1.4 Выбор пространственной матрицы с помощью имитационного моделирования 41

2.2 Пространственный эконометрический анализ региональной безработицы 47

2.2.1 Факторы, объясняющие региональные различия в уровнях безработицы 49

2.2.2 Пространственная эконометрическая модель региональной безработицы в России 55

2.2.3 Пространственная эконометрическая модель региональной безработицы в Германии 74

**Глава 3 Прогнозирование региональной безработицы с помощью пространственных эконометрических моделей 96**

3.1 Обзор исследований, использующих пространственные модели для прогнозирования на региональных данных 98

3.2 Модели, используемые для предсказания

3.2.1 Основные виды моделей 102

3.2.2 Объясняющие переменные 103

3.2.3 Оценка моделей для прогнозирования 106

3. 3 Прогнозирование уровня безработицы 107

3.3.1 Формулы для прогнозов 107

3.3.2 Меры для сравнения качества прогнозов 109

3.3.3 Результаты прогнозирования 110

Заключение 115

Список литературы

* [Обзор эмпирических исследований](http://www.dslib.net/mat-metody/ispolzovanie-prostranstvennoj-zavisimosti-v-regionalnyh-issledovanijah-na.html#7665171)
* [Выбор пространственной матрицы с помощью имитационного моделирования](http://www.dslib.net/mat-metody/ispolzovanie-prostranstvennoj-zavisimosti-v-regionalnyh-issledovanijah-na.html#7665172)
* [Модели, используемые для предсказания](http://www.dslib.net/mat-metody/ispolzovanie-prostranstvennoj-zavisimosti-v-regionalnyh-issledovanijah-na.html#7665173)
* [Меры для сравнения качества прогнозов](http://www.dslib.net/mat-metody/ispolzovanie-prostranstvennoj-zavisimosti-v-regionalnyh-issledovanijah-na.html#7665174)

## Обзор эмпирических исследований

Развитие идей пространственной эконометрики тесно связано с зарождением идеи о пространственной автокорреляции. Среди первых упоминаний об этой концепции Getis [2008] отмечает Эрнеста Георга Равенштейна, который писал о миграции в 1885, а также работу Тюнена «Изолированное государство» [1826]. Активно концепция пространственной автокорреляции начала развиваться в 50-х годах XX века в Вашингтонском университете благодаря работам M. F. Dacey, W. L. Garisson, E. Ulmann и других. Перечисленные географы отмечали важным взаимное расположение объектов относительно друг друга. Именно оно, по их мнению, играло наиболее значимую роль для объяснения природы человеческой активности.

Существующие на тот момент коэффициенты пространственной корреляции назывались коэффициентами смежности (сontiguity ratios) [см. Geary, 1954; Moran, 1948]. Geary обращал внимание на то, что условие независимости остатков регрессии часто не выполняется из-за взаимного влияния наблюдений, расположенных близко друг к другу. Позже идею учета пространственной корреляции в уравнении регрессии выдвинули Thomas и др. [1965], продолжая идею Робинсона [Robinson, 1956] о необходимости взвешивать наблюдения в соответствии с влиянием каждого наблюдения.

Особое влияние на развитие пространственной эконометрики оказала книга «The Problem of Spatial Autocorrelation», положившая начало региональным исследованиям. Cliff и др. [1968] обратили внимание на проблему спецификации моделей в пространственном анализе, а именно на неверную спецификацию моделей, не учитывающих пространственную автокорреляцию. Позже авторы опубликовали книгу «Spatial Autocorrelation» [Cliff и др., 1973], в которой продемонстрировали, как в регрессионном анализе остатки могут тестироваться на присутствие пространственной автокорреляции.

Рождение пространственной эконометрики как самостоятельной дисциплины обычно связывают с возникновением термина «пространственная эконометрика». В мае 1974 на Ежегодной встрече Датской Статистической Ассоциации в Тильбурге Paelink заявил о необходимости основать новую отрасль эконометрики, обеспечивающую методологические основы для региональных и городских эконометрических моделей, и дал ей название «пространственная эконометрика». Вместе с Klaassen в первой книге по пространственной эконометрике [Paelinck и др., 1979] они предложили тесты на выявление значимости пространственных параметров, которые и по сей день играют ключевую роль в данной дисциплине.

Следующий фундаментальный вклад в развитие пространственной эконометрики внес Anselin в 1988 г. В его книге освещаются основные принципы пространственной корреляции, ее учета в регрессионных моделях, а также последствия неправильной спецификации. В своей книге автор определяет пространственную эконометрику как «набор методов, которые работают с особенностями, вызванными пространственным расположением, в моделях регионального анализа» (с. 7).

Несмотря на то, что к концу 80-х основные пространственно-эконометрические модели и методы их оценки уже были предложены в литературе, они еще не представляли интерес для широкого круга экономистов, в то время как методы анализа временных рядов были обязательной частью учебников по эконометрике. Однако было несколько работ, которые говорили о необходимости учитывать пространственную корреляцию. В частности, Arbia [2006] указывает на следующих авторов. Например, Johnston [1991] отмечает, что аналогично корреляции остатков во временных рядах, пространственная автокорреляция может проявляться и в случае с кросс-секционными данными. Kmenta [1986] указывает на то, что предположение о независимости остатков часто подвергается сомнениям: это предположение может не выполняться в случае, например, если наблюдениями являются регионы. Упоминание о пространственной корреляции также встречается в книгах Baltagi [2001]; Maddala [2001]. Kennedy [2003] наряду с упоминанием об автокорреляции остатков во временных рядах говорит о возможном существовании автокорреляции остатков в пространстве.

Основополагающую роль в развитии пространственной экономики, а следовательно и пространственной эконометрики, сыграла идея Диксита-Стиглица, подстегнувшая интерес к несовершенной конкуренции и возрастающей отдаче от масштаба. Развитие на основе этих идей Новой Экономической Географии, предложенных Кругманом, стало дополнительным стимулом развития пространственной эконометрики. Немалую роль для развития методов сыграла также возросшая популярность анализа эффектов соседства в социологии и развитие теории социального взаимодействия.

## Выбор пространственной матрицы с помощью имитационного моделирования

Анализ рынка труда в региональном разрезе является необходимым дополнением моделирования национального рынка труда. В связи с тем, что решения участников рынков труда, например, решение о миграции с целью поиска работы в другом регионе, не ограничиваются границами одного региона, региональные рынки труда зависят от взаимного расположения соответствующих регионов. Локальный шок на рынке труда одного региона может отражаться и на других регионах. Таким образом, национальный рынок труда является системой связанных между собой территориальных рынков [Ощепков и др., 2015]. Важность расширения исследований региональных российских рынков труда отмечена в работе [Блинова и др., 2002]. Результаты их работы подтверждают, что для создания «эффективной структуры занятости» унифицированный подход, в котором рынок труда анализируется в целом по России, малоэффективен, и необходимо «учитывать отраслевую специализацию региональных рынков труда и факторов экономической политики на уровне региона».

Рынок труда динамичен не только во времени, но и в пространстве. Это означает, что изменение показателей рынка труда в одном регионе повлечет за собой изменения на рынках труда других регионов. Потоки трудовой мобильности (миграционные потоки, обусловленные поиском работы; маятниковая трудовая миграция) тесно связывают между собой региональные рынки труда, которые образуют единый общенациональный рынок. Вследствие того, что природа процессов неразрывно связана с географическим взаиморасположением регионов, необходимость учитывать пространственное взаимодействие, вызванное потоками трудовой мобильности, не вызывает сомнений.

Методология пространственного эконометрического моделирования позволяет учитывать географическую структуру данных путем включения пространственного лага (взвешенного значения показателя в других регионах) в модель регрессии. При анализе региональной безработицы, например, ключевым фактором, который необходимо учитывать в моделях, является взвешенный уровень безработицы других регионов. Обычно чем ближе регионы расположены друг к другу, тем активнее они влияют друг на друга. Описываемый эконометрический подход позволяет учитывать пространственную корреляцию между объектами и их взаимное влияние друг на друга (spillover effect). Например, наличие положительного пространственного эффекта означает, что увеличение или снижение уровня безработицы в одном регионе повлечет за собой соответственно увеличение или снижение уровня безработицы в соседних регионах. Игнорирование географической структуры при анализе объектов с постоянным местоположением влечет за собой определенные методологические проблемы. Например, если в модели не учитывать показатели соседних объектов, то оценки, полученные методом наименьших квадратов или методом максимального правдоподобия, будут смещенными и несостоятельными [Cliff, Ord, 1981]. В связи с вышесказанным появилось большое число исследований, в которых показатели занятости изучаются при помощи пространственных-эконометрических моделей.

Необходимость моделирования взаимосвязей между уровнем безработицы и характеристиками рынков труда на региональном уровне [Elhorst, 2003] объясняется рядом причин. Во-первых, причины региональных различий отличаются от причин межстрановых различий в показателях рынка труда. Различия в показателях рынков труда между странами в основном обусловлены различными институтами, например, такими как пенсионная и налоговая системы, программы социальной политики и т.п. Однако в России, как и во многих других странах, эти институты едины для всей страны, следовательно, для объяснений межрегиональных различий требуются другие источники. Во-вторых, степень различий между показателями на региональном уровне не меньше, чем на межстрановом. Степень межрегиональных различий в уровнях безработицы велика: в регионах России за 2005–2012 гг. уровень безработицы принимает значения от 0,8% до 23,9% (см. табл. Б.1 в Приложении раздела 2.2.2); в регионах Германии за 2001–2009 гг. уровень безработицы принимает значения от 1,96% до 31,14% (см. табл. В.1 в Приложении раздела 2.2.3). В-третьих, неравенство в уровнях занятости и безработицы может быть неэффективным. Поэтому снижение этого неравенства уменьшает негативные последствия от высокой концентрации безработицы в определенных регионах и может приводить к таким положительным последствиям, как более высокий экономический рост и снижение инфляции [Elhorst, 2003; Taylor, 1996]. Россия, как и Германия, отличаются высокой региональной дифференциацией в том числе и по уровню безработицы [Guriev и др., 2012], поэтому важно изучать рынок труда и на уровне регионов.

Данный раздел второй главы посвящен пространственному эконометрическому анализу региональной безработицы в России и Германии. В разделе 2.2.1 приводится описание факторов, влияющих на региональный уровень безработицы, раздел 2.2.2 посвящен анализу региональной безработицы в России, а в разделе 2.2.3 представлен анализ региональной безработицы в Германии.

В текущем диссертационном исследовании при выборе факторов, объясняющих уровень безработицы, используется подход, предложенный в работах [Marston, 1985; Partridge и др., 1995]. В соответствии с этим подходом в оцениваемом уравнении все показатели можно разделить на две основных группы: показатели, объясняющие безработицу с точки зрения теории о равновесном уровне безработицы, и показатели, объясняющие безработицу с точки зрения теории о неравновесном уровне безработицы.

Теория о неравновесном уровне безработицы предполагает, что уровень безработицы достигает своего равновесного значения только в долгосрочном периоде вследствие медленной скорости подстраивания. Поэтому различия в уровнях безработицы сохраняются на протяжении долгого периода времени. Основным определяющим фактором для региональной безработицы в рамках данной теории является скорость установления равновесия. Поэтому существенно важными здесь представляются миграционные процессы. Скорость приспособления в свою очередь определяется спросом и предложением на труд. В частности, она зависит от гибкости заработных плат, решений домохозяйств о миграции и участии в рабочей силе, решений фирм о местоположении [Aragon, Haughton, Haughton, Leconte, Malin, Ruiz-Gazen, Thomas-Agnan, 2003].

## Модели, используемые для предсказания

Напомним, что существует три различных способа оценки динамических пространственных моделей панельных данных: с помощью обобщенного метода моментов, метода максимального правдоподобия, метода Монте-Карло на марковских цепях (MCMC). В данной главе применяется два метода оценивания: метод максимального правдоподобия (ММП) для оценки моделей (2.23) и (2.24) и обобщенный метод моментов (IV/GMM) для оценки моделей (2.24), (2.25) и (2.26). Также применяют байесовский подход, оценивая модели с помощью метода Монте-Карло по схеме марковской цепи (MCMC) (LeSage, 1997).

Метод максимального правдоподобия в применении к динамическим пространственным моделям был рассмотрен в работе [Yu и др., 2008], а в работе [Lee, Yu, 2010] была предложена его коррекция для устранения проблемы несостоятельности оценок, о которой говорилось выше. Его применение корректно, когда либо Yt-1, либо WYt-1 не включено в уравнение регрессии [Elhorst, 2012].

Далее для оценивания динамической модели применяется IV/GMM подход, основанный на методе Ареллано–Бонда [Arellano и др., 1991; Blundell и др., 1998]. Проблема использования стандартного обобщенного метода моментов (difference GMM) состоит в том, что данная техника оценивания приводит к смещенным оценкам, что в особенности касается коэффициента [Elhorst, 2010]. Lee, Yu [2010] показали наличие смещения в оценках и предложили модификацию обобщенного метода моментов, которая дает состоятельные оценки даже при малом T и большом N. Другой способ получить состоятельные оценки, основываясь на обобщенном методе моментов, заключается в использовании ОММ [Monteiro и др., 2009]. Monteiro, Kukenova [2009] приходят к выводу, что смещение в оценках, показанное в работе [Lee, Yu, 2010], существенно уменьшается в случае использования ОММ и даже может быть проигнорировано. Использование метода IV/GMM мотивировано тем, что он справляется с лаговой эндогенностью, возникающей при элиминировании фиксированных эффектов первыми разностями. Кроме того, использование данного метода необходимо в связи с эндогенностью некоторых регрессоров, например, заработной платы или числа ежедневно мигрирующих. В качестве инструментов берутся лаги переменных модели.

В настоящей работе применяются оба метода. В связи с наличием корреляции второго порядка ошибок в первых разностях, которая была выявлена с помощью теста Aреллано-Бонда, в модель, оцениваемую с помощью ОММ, включен второй лаг. В эту модель дополнительно включены следующие переменные: число ежедневно мигрирующих и число вновь зарегистрированных фирм. Эти переменные не могли быть включены в модель, оцениваемую ММП, ввиду недостатка некоторых значений.

Зависимость Y от своего пространственного лага WY не обязательно линейна. На основе диаграммы рассеяния выдвинуто предположение, что квадратичная функция более адекватно описывает эту зависимость, чем линейная функция. Поэтому для пространственного лага введена квадратичная зависимость. Модель, оцениваемая с помощью ОММ, принимает следующий вид: У и = ВД-,,-1 + Т2У,, -2 + PWy» + P2Wy2" + x t P + V,+ Pt+ v., (2.25) где p2Wy2it обозначает квадрат взвешенной суммы уровней безработицы остальных регионов. Модель оценивается с помощью метода Бланделла-Бонда [Blundell, Bond, 1998]. Оценка пространственных эффектов между Востоком и Западом

Логично предположить, что Западная и Восточная Германия могут иметь различия как во влиянии рассматриваемых факторов на уровень безработицы, так и в пространственных эффектах. Lottmann [2012] анализирует различия между пространственными эффектами Западной и Восточной Германии путем оценивания моделей на подвыборках. Такой подход дает возможность оценить различие между влиянием экзогенных переменных на уровень безработицы в разных частях страны. В ее работе было показано, что большинство коэффициентов значимо различаются в моделях для восточных и западных регионов.

Однако данный подход не позволяет выявить пространственные эффекты влияния восточных регионов на западные и наоборот. Для выявления таких эффектов применяется подход, который был разработан для российских регионов [Демидова, 2014]. Используется следующее расширение спецификации: где yi wt и yiet — объясняемые переменные западных и восточных регионов соответственно, Xtw и Xte — объясняющие переменные западных и восточных регионов соответственно, t — временные эффекты, i — индивидуальные эффекты, it — вектор остатков.

## Меры для сравнения качества прогнозов

Как описано выше, прогнозы построены для двух случаев: когда объясняющие переменные в будущем периоде известны, и имеется возможность строить прогнозы в соответствующих точках, и когда объясняющие переменные в будущем периоде неизвестны, и необходимо использовать их лаги.

В первом случае оцениваются модели 1-11 для случая, когда значения соответствующих регрессоров известны. Результаты оценки качества прогнозов для российских регионов на один год вперед представлены в таблице Г.1 (см. Приложение главы 3). Для оценки этих моделей использовались данные за 2005-2010 гг. (по России и Германии). В случае предсказания уровня безработицы в регионах России можно заметить, что учет пространственных эффектов значительно улучшает качество предсказания: c RMSE = \A% до RMSE = 0.99 . Включение динамического лага в модель незначительно улучшает качество прогноза.

Интересно отметить, что наивные предсказания обладают достаточно неплохой прогнозной силой: качество прогноза, вычисленного как среднее предыдущих значений, сравнимо с прогнозом, полученным с помощью модели с фиксированными эффектами без учета пространственных эффектов ( RMSE =1.32 против RMSE =1.48 ). Разделение данных на восточные и западные регионы значительно улучшает качество предсказания для восточных регионов (с RMSE =0.99 до RMSE =0.54).

В случае предсказания безработицы в Германии (см. таблицу Г.3 Приложения главы 3) пространственная динамическая модель с фиксированными эффектами дает наилучшие результаты ( RMSE =0.45). Качество прогнозов этой модели значительно превосходит качество наивных прогнозов. Добавление пространственных эффектов в модель с фиксированными эффектами улучшает прогноз (с RMSE =1.54 до RMSE =1.28) в случае использования граничной матрицы. При оценке моделей на подвыборках для востока и запада наилучший прогноз для западных регионов получается при использовании пространственной динамической модели с фиксированными эффектами без объясняющих переменных, а для восточных регионов — пространственной динамической модели с фиксированными эффектами.

Результаты оценки качества предсказания на два шага вперед для российских регионов11 (см. таблицу Г.2) подтверждают, что включение пространственных эффектов в модель улучшает прогноз (с RMSE =2.43 до RMSE =1.48). Наилучшее качество прогноза по-прежнему достигается пространственной моделью с фиксированными эффектами и матрицей обратных расстояний. При этом разделение выборки на восточные и западные регионы улучшает качество предсказания.

Таким образом, если известны значения соответствующих регрессоров, либо их значения могут быть хорошо спрогнозированы, то для прогнозирования безработицы на региональном уровне имеет смысл использовать пространственные модели панельных данных: качество их прогноза является наилучшим из представленных моделей, в том числе по сравнению с часто используемыми наивными прогнозами.

К сожалению, не всегда имеется возможность качественно спрогнозировать объясняющие переменные. В этом случае возможным выходом из положения может стать использование лагов регрессоров. Тогда проверяется качество прогнозов, полученных с помощью моделей 1– 11, где текущие значения регрессоров заменены на значения из предыдущего периода.

В таблицах Г.4 и Г.5 (в приложении главы 3) представлены результаты качества предсказаний на 2013 г. (для России) и 2012 г. (для Германии). Для оценки данных моделей использовались данные за 2006– 2012 гг. (по России) и 2006–2011 гг. (по Германии). Учет

пространственных взаимосвязей по-прежнему позволяет получить более качественный прогноз по сравнению с классическими моделями панельных данных. Однако наилучшим прогнозом в данном случае является наивный прогноз: в среднем по всем регионам его качество немного превосходит качество прогноза, полученного с помощью динамической пространственной модели ( RMSE =0.72 против RMSE =0.73 для России и RMSE =0.34 против RMSE =0.39 для Германии).

Тем не менее, динамическая пространственная модель все еще остается более предпочтительной для некоторых отдельных регионов. В целях проверки робастности, качество прогнозов на один шаг вперед, полученных с помощью моделей 1–11, где текущие значения регрессоров заменены на предыдущие значения, было проверено для других периодов. Были построены прогнозы оценки моделей использовались следующие соответствующие временные промежутки: 2006–2011 гг., 2006–2010 гг., 2006–2009 гг. (для оценки данных по России) и 2006–2010 гг., 2006–2009 гг. (для оценки данных по Германиина 2012, 2011 и 2010 гг. (для России) и 2011, 2010 гг. (для Германии). Для). В приложении главы 3 приведен список регионов, для которых во всех временных периодах качество прогноза по пространственной модели превосходит качество прогноза, полученного с помощью наивной модели (12 регионов для России и 28 регионов для Германии).

Включение пространственных эффектов позволяет получить лучшее качество предсказания безработицы в региональном разрезе в сравнении с классическими моделями панельных данных. Настоящее исследование дополнительно отвечает и на вопрос, имеет ли смысл использовать сложные пространственные модели панельных данных вместо простых в построении наивных предсказаний, часто используемых в практике? Если значения объясняющих переменных известны или имеются их предсказанные значения, то для построения прогноза зависимой переменной стоит использовать пространственные модели. Если значения объясняющих переменных неизвестны, и в моделях используются их лаги, то прогнозы, полученные с помощью пространственных моделей, немного уступают наивному предсказанию по среднему показателю качества, тем не менее, для некоторых отдельных регионов они все же позволяют улучшить качество прогноза.

Таким образом, в данной главе осуществлена попытка прогнозирования уровня безработицы в регионах России и Германии на основе пространственных моделей панельных данных. Показано, как учет пространственных эффектов в модели влияет на качество прогноза, а также получен ответ на вопрос, позволяют ли пространственные эконометрические модели получить лучшее качество прогноза, чем наивные прогнозы.