Прикладная эконометрика, 2016, т. 43, с. 29–51.
Applied Econometrics, 2016, v. 43, pp. 29–51.

Е. В. Семерикова, О. А. Демидова[1]

Использование пространственных эконометрических моделей при прогнозе регионального уровня безработицы

Рассматривается вопрос прогнозирования уровня безработицы в регионах России и Германии с помощью эконометрических моделей панельных данных. На региональных данных за 2005–2012 гг. показано, что прогнозы пространственных моделей панельных данных превосходят качество прогнозов других моделей (в среднем по всем регионам или для некоторых из них): непространственных моделей панельных данных, моделей сквозной регрессии, моделей без учета объясняющих переменных, а также наивных прогнозов (среднее значение за один или несколько предыдущих периодов).

Ключевые слова: пространственные модели панельных данных; региональная безработица; прогноз; регионы России, регионы Германии.

JEL classification: R1; C21; C23.

1. Введение

Прогнозирование основных показателей состояния рынков труда имеет большое экономическое и социальное значение для многих стран, в том числе для России. Например, в рамках Гайдаровского форума 15 января 2016 г. состоялась экспертная дискуссия «Проблемы прогнозирования и моделирования рынка труда в России». Одним из самых важных показателей рынка труда является уровень безработицы, прогнозированию которого посвящено множество работ на данных различных стран. Техника прогнозирования достаточно хорошо развита для временных рядов, поэтому не удивительно, что исследователи в подавляющем большинстве случаев используют данные этого типа. Однако нередко возникает дилемма, какой частоты данные надо использовать. Ряд с промежутком в один год может оказаться недостаточно длинным, а при использовании квартальных и месячных данных на первый план выходит проблема сезонности (затрудняющая применение традиционных техник исследования временных рядов с помощью методологии Бокса–Дженкинса и поиска коинтеграционных соотношений).

В связи с этим все чаще появляются исследования, авторы которых (далее мы остановимся на некоторых исследованиях) используют данные не для стран, а для более мелких

[1] **Семерикова Елена Вячеславовна —** Национальный исследовательский университет «Высшая школа экономики», Москва; lena.sem7@mail.ru.

Демидова Ольга Анатольевна — Национальный исследовательский университет «Высшая школа экономики», Москва; demidova@hse.ru.

административных единиц — регионов. Это позволяет использовать панельные данные, получать более детальные прогнозы, что представляет интерес для стран со значительным региональным разнообразием, например, России и Германии. Предсказание регионального уровня безработицы не менее важно, чем предсказание безработицы для страны в целом. Это может оказать помощь при разработке внутренней экономической политики, к примеру, как распределить ресурсы и расходы между регионами. Показатели рынков труда зачастую имеют более существенные различия между регионами, чем между странами (Overman et al., 2002), и уровень региональной безработицы может существенно отличаться от общенационального. Предсказание региональных показателей является достаточно сложной задачей, в том числе и потому, что регионы взаимосвязаны, и необходимо учитывать их взаимное влияние друг на друга.

В настоящей работе осуществлена попытка предсказать уровень безработицы в региональном разрезе для России и Германии. С помощью панельных данных учитывается региональная гетерогенность, а используемый в работе пространственно-эконометрический подход позволяет учесть взаимное влияние регионов друг на друга. Располагая данными за 2005–2012 гг. (для России) и 2005–2011 гг. (для Германии), авторы строят предсказания для 2011–2012 гг. в предположении, что объясняющие переменные известны. Также строится прогноз на дополнительный год в предположении, что объясняющие переменные неизвестны, вследствие чего используются их лаги в оцениваемой модели.

Основная цель работы — проверить, насколько включение пространственных эффектов улучшает качество прогноза. Прогноз, полученный с помощью пространственных эконометрических моделей, сравнивается с предсказаниями на основе других распространенных моделей панельных данных. Кроме того, сравнивается качество предсказания с наивными прогнозами, рассчитанными как предыдущее значение уровня безработицы и среднее значение предыдущих значений за несколько лет. В качестве критериев сравнения используются средняя квадратичная ошибка (RMSE), средняя абсолютная ошибка в процентах (MAPE), средняя абсолютная ошибка (MAE), симметричная абсолютная ошибка в процентах (sMAPE).

В результате построения и сравнения прогнозов выявлено, что в случае известных значений объясняющих переменных пространственные модели позволяют получить наилучшее качество прогноза. В случае использования лагов объясняющих переменных пространственные модели в среднем незначительно «проигрывают» наивным предсказаниям, однако для некоторых регионов позволяют получить лучшие предсказания. Таким образом, учет пространственных эффектов в модели позволяет получить высокое качество предсказаний.

В следующем разделе представлен обзор литературы, посвященной вопросу прогнозирования региональных показателей. Далее описываются спецификации оцениваемых моделей, используемые данные и переменные, а также методы оценивания моделей. Четвертый раздел статьи посвящен прогнозированию: описаны формулы предсказаний и меры измерения их качества, представлены результаты. Пятый раздел содержит заключительные выводы.

2. Обзор литературы

Большая часть исследований, в которых рассматривается вопрос прогнозирования на региональном уровне, описаны в работе Lehmann, Wohlrabe (2014). Авторы отмечают, что прогнозирование с использованием региональных данных становится все более популярным,

причем в большинстве исследований основное внимание направлено на предсказание валового регионального продукта (ВРП) или показателей рынка труда, таких как безработица или занятость. Mayor, Patuelli (2012) показывают, что в случае длинных временных промежутков модели временных рядов позволяют получить лучшее качество прогноза, чем пространственные модели, однако в случае небольшого числа наблюдений во времени «выигрывают» пространственные модели.

Остановимся подробнее на двух типах используемых моделей. В первом типе моделей участвует только зависимая переменная с временными и пространственными лагами (Schanne et al., 2010), а во втором — дополнительно включены независимые объясняющие переменные.

Преимуществом моделей первого типа является простота построения прогноза. Аналогично прогнозам для временных рядов, для этих моделей требуются лишь значения зависимой переменной на одном или нескольких предыдущих шагах, а включение пространственных лагов не сильно усложняет схему (подробнее см. в разделе 3.1). Schanne et al. (2010), используя месячные данные о региональных показателях рынка труда (в том числе уровне безработицы) для 176 немецких регионов с января 1998 г. по декабрь 2007 г., сравнивают ARIMA и GVAR модели (часть наблюдений использовалась для оценки параметров моделей, а остальные — для проверки качества прогноза на основе соответствующей модели). Было показано, что модели GVAR, включающие пространственные лаги, позволяют сделать более точные прогнозы на 12 месяцев 2004, 2005, 2006, 2007 гг. Kholodilin et al. (2008), используя данные для 16 немецких земель за 1993–2001 гг., сравнивают прогнозы для ВРП на 2002–2006 гг., полученные с помощью динамических моделей панельных данных с независимыми ошибками, и прогнозы, полученные с помощью динамических моделей с пространственно-коррелированными ошибками (SAE, spatially autocorrelated errors). Последние модели позволяют улучшить качество предсказания (в терминах показателя RMSE) на 25%. Авторы обоих исследований отмечают, что включение в модель пространственных эффектов позволяет лучше учесть разницу, существующую между восточными и западными немецкими регионами. В другом исследовании при построении многошагового прогноза ежегодных темпов роста ВРП в 31 китайской провинции Girardin, Kholodilin (2011) также сравнивают прогнозы, полученные с помощью динамических моделей с включением пространственных лагов и без них (взвешенного роста темпа ВРП в соседних провинциях). Пространственно-эконометрические модели позволяют получить более точное предсказание на каждом шаге, причем «выгода» от учета пространственной структуры увеличивается в случае более длинных горизонтов прогнозирования: при прогнозировании на один год точность прогноза увеличивается на 8% (в терминах средней квадратичной ошибки), а при прогнозировании на 14 лет — более чем на 25%.

Модели второго типа дают возможность учесть дополнительную информацию, содержащуюся в независимых переменных, что позволяет надеяться на более точный прогноз. Но в этом случае возникает проблема, что при оценке модели на некотором временном интервале для прогноза на один или несколько шагов необходимы будущие значения объясняющих переменных, а они недоступны. Поэтому приходится либо использовать лаги независимых переменных (X_{t-1} и т. п. вместо X_t), что сделано, например, в работе Longhi, Nijkamp (2007), либо оценивать систему уравнений (Polasek et al., 2007), либо использовать имитационное моделирование данных (Baltagi, Li, 2006), либо строить прогнозы для неза-

висимых переменных (Ачкасов, Шатило, 2016). Приведем дополнительные подробности перечисленных статей.

Baltagi et al. (2014), используя данные имитационного моделирования, сравнивают прогнозы по динамическим моделям, оцененным с помощью обобщенного метода моментов. Полученные результаты свидетельствуют о преимуществе прогнозирования на основе моделей, в которые включены пространственные эффекты. Для данного исследования этот результат не удивителен, поскольку процесс, с помощью которого имитировались данные, имеет авторегрессионную природу. Однако этот же результат нередко имел место и для реальных данных. Longhi, Nijkamp (2007) прогнозируют занятость в 326 западных регионах Германии в 2000–2003 гг. по данным за 13 предыдущих лет. В качестве объясняющих переменных используются первые лаги занятости в различных секторах экономики. Прогнозы этой простейшей модели существенно улучшаются при добавлении пространственно-временного лага переменной «заработная плата» или автокоррелированной структуры ошибок. Polasek et al. (2007) по данным для 99 австрийских регионов за 1998–2003 гг. продемонстрировали, как с помощью динамической системы уравнений с пространственными лагами и зависимыми переменными (такими как ВРП, плотность населения, уровень занятости) можно построить прогноз на 20 лет. Для оценки параметров были использованы байесовские методы.

Перечисленный выше список работ, в которых прогнозы делаются на основе моделей, оцененных по региональным данным с включением пространственных эффектов, не является полным. Авторы не ставили перед собой задачу охватить все исследования, а хотели лишь кратко описать идеи, лежащие в основе этого типа прогнозирования. Но даже упомянутые работы дают представление о том, что исследований с использованием немецких региональных данных немало, а по российским регионам почти нет. Отметим, что для прогнозов уровня безработицы в России обычно используются данные по России в целом. Прогнозы часто строятся с помощью векторной модели коррекции ошибок (Ачкасов, Шатило, 2016; Гурвич, Вакуленко, 2016), моделей Бокса–Дженкинса (Турунцева, Астафьева, 2016) и т. п.

В данном исследовании авторы хотели частично восполнить пробел, связанный с отсутствием прогнозов для региональной безработицы в России, сравнить разные типы моделей, в том числе модели с использованием пространственных эффектов и без них. В качестве последних используются не только модели без объясняющих факторов, но и совсем простые, наивные модели, в которых прогноз вычисляется как среднее за один или несколько предшествующих периодов. По имеющимся сведениям, пространственные прогнозы для Германии не сравнивались с наивными, к тому же авторам не известны исследования, в которых использовались данные за последние годы. В связи с этим было решено сравнить качество прогнозов по региональным данным не только для России, но и для Германии.

В следующем разделе дано описание моделей, которые используются для предсказания. Среди них перечислены упомянутые выше наивные модели, модели первого типа, включающие только временные и пространственные лаги зависимой переменной, и модели второго типа, включающие независимые переменные. При построении прогнозов по моделям второго типа рассматриваются два случая: 1) объясняющие переменные в будущем периоде известны, и можно строить прогнозы в соответствующих точках; 2) объясняющие переменные в будущем периоде неизвестны и используются их лаги.

3. Модели, используемые для предсказания

3.1. Основные виды моделей

В данной работе за основу взята самая общая динамическая модель с индивидуальными и временными эффектами, а также пространственным лагом (SAR):

$$y_{it} = \rho W y_{it} + \gamma y_{it-1} + X'_{it}\beta + \mu_i + \varphi_t + \varepsilon_{it}, \quad (1)$$

где $i = 1,\ldots,N$ — номер региона (для России $N = 75$, для Германии $N = 370$),
$t = 2005,\ldots,T$ (для разных моделей T варьируется от 2010 до 2012),
Y_{it} — уровень безработицы в регионе i в момент t,
Wy_{it} — пространственный лаг зависимой переменной,
y_{it-1} — динамический лаг зависимой переменной,
X_{it} — набор объясняющих переменных (описанных в следующем разделе),
μ_i — индивидуальные эффекты,
φ_t — временные эффекты.

В случае оценки моделей с фиксированными эффектами $\varepsilon_{it} \sim N(0,\sigma_i^2)$ (для стандартных отклонений оценок коэффициентов в этом случае использовались робастные оценки в форме Уайта). В случае оценки моделей со случайными эффектами предполагалось, что дисперсии ошибок одинаковы: $\varepsilon_{it} \sim N(0,\sigma^2)$.

Были рассмотрены также частные случаи общей модели:

1) модель сквозной регрессии ($\rho = \gamma = 0$, $\mu_i = \mu \ \ \forall i$);

2) модель без пространственных эффектов ($\rho = \gamma = 0$) с фиксированными эффектами (μ_i — неизвестные параметры);

3) модель без пространственных эффектов ($\rho = \gamma = 0$) со случайными эффектами ($\mu_i = \mu + u_i$, где μ — параметр, общий для всех регионов, а u_i — ошибки, некоррелированные с ε_{it} и некоррелированные при разных i);

4) пространственная модель ($\gamma = 0$) с фиксированными эффектами;

5) пространственная модель ($\gamma = 0$) со случайными эффектами;

6) динамическая модель без объясняющих переменных ($\beta = 0$) с фиксированными эффектами;

7) динамическая модель без объясняющих переменных ($\beta = 0$) со случайными эффектами;

8) модель без объясняющих переменных и динамического лага ($\beta = \gamma = 0$) с фиксированными эффектами;

9) модель без объясняющих переменных и динамического лага ($\beta = \gamma = 0$) со случайными эффектами;

10) модель без пространственных эффектов ($\rho = \gamma = 0$) с фиксированными эффектами;

11) модель без пространственных эффектов ($\rho = \gamma = 0$) со случайными эффектами.

К вышеперечисленным моделям были также добавлены их версии без учета временных эффектов: $\phi_t = 0 \ \ \forall t$.

3.2. Объясняющие переменные

Для оценки моделей использовались панельные данные о 75 регионах России[2] и 370 регионах Германии за период 2005–2012 гг. Данные о российских регионах обеспечены Федеральной службой государственной статистики (Регионы России, 2005–2012), данные о регионах Германии — Статистическим бюро Германии (www.regionalstatistik.de).

В качестве объясняющих факторов уровня безработицы используют различные экономические и демографические показатели. Основной фактор, отражающий экономическую ситуацию в регионе — валовый региональный продукт. Это наиболее часто используемый индикатор регионального спроса на труд. Определяющим фактором является также отраслевая структура экономики (Armstrong, Taylor, 1988; Elhorst, 2003;, Martin, 1997), отражающая различия спроса на труд между регионами, поэтому в состав объясняющих переменных включены доли занятых в различных отраслях (подробнее см. табл. 1). На спрос и предложение труда оказывает влияние в том числе и уровень миграции (Elhorst, 2003), поэтому этот фактор тоже учитывается в текущем исследовании. Также важную роль играют издержки миграции, являющиеся определенным регулятором миграционных потоков между регионами. В качестве индикаторов издержек миграции используют, например, показатели стоимости и доступности жилья в регионе (Aragon et al., 2003).

Необходимость учета заработной платы подтверждается наличием большого числа работ, где обсуждается влияние заработной платы на безработицу (например, (Molho, 1995; Partridge, Rickman, 1995)) и чаще всего выявляется их отрицательная взаимосвязь.

Если рынки труда не находятся в равновесии, отклонения уровня безработицы от среднего уровня могут быть вызваны изменениями предложения труда, технологическими или иными шоками (Martin, 1997, p. 244). Поэтому в моделях также важно учитывать такие факторы, как возрастная структура населения и уровень образования (Aragon et al., 2003; Hofler, Murphy, 1989). Кроме того, на скорость реагирования рынков на внешние шоки влияет уровень урбанизации. В нашей модели в качестве соответствующей переменной выбрана доля городского населения.

Еще одним фактором, определяющим уровень безработицы, является плотность населения. С одной стороны, высокая плотность населения изменяет скорость подстройки рынка (Elhorst, 2003), с другой стороны, она может отражать степень привлекательности региона. Как и в работе (Lottmann, 2012), посвященной региональной безработице в Германии, в настоящем исследовании для оценки безработицы в Германии также учитывается рост занятости. Кроме того, среди определяющих факторов часто используется уровень преступности (Greenwood et al., 1991; Marston, 1985), отражающий экономическую и социальную ситуацию в регионе.

Таким образом, в качестве факторов для России и Германии были выбраны отраслевая структура (например, доля занятых в сельскохозяйственной отрасли в общем числе занятых[3]), возрастной состав занятого населения, уровень образования занятых, коэффициент

[2] Данные о некоторых регионах России не включены в исследование. Так, ввиду административных реформ не включен Забайкальский край (объединение Читинской области и Агинского Бурятского АО, 01.03.2008). Ввиду отсутствия данных по некоторым переменным не были включены: Чеченская республика, Чукотский АО, Республика Ингушетия, Республика Дагестан, а также Ненецкий АО (в составе Архангельской области), Ханты-Мансийский и Ямало-Ненецкий АО (в составе Тюменской области).

[3] Полный список отраслей представлен в табл. 1.

миграции, валовый региональный продукт, плотность населения и средний уровень заработной платы в регионе. Дополнительно к этим переменным для России были включены доля городского населения, доступность жилья и уровень преступности. В базу данных по Германии дополнительно включен темп роста занятого населения в регионе. Незначительное различие наборов объясняющих переменных между базами данных обусловлено доступностью тех или иных данных для выбранного временного промежутка.

Данные о валовом региональном продукте на душу населения в базовых ценах скорректированы на покупательную способность путем деления на ИПЦ (для России и Германии) и стоимость потребительской корзины для каждого региона (для России). В качестве переменных, характеризующих возрастную структуру и образование населения, взяты доли занятых в определенных возрастных категориях, а также доли занятых с высшим, профессиональным и школьным образованием. В качестве показателя издержек миграции в текущей работе использовался показатель доступности жилья, рассчитанный как средняя в регионе цена за 1 кв. м жилья на вторичном рынке, скорректированная на средний доход домохозяйств в регионе.

Полный перечень объясняющих переменных представлен в табл. 1. После оценки некоторые незначимые объясняющие переменные были исключены из модели. Исключение переменных проводилось последовательно, с предварительной проверкой гипотезы о равенстве нулю коэффициента перед переменной, а также проверкой на улучшение качества предсказания после удаления переменной. Часть незначимых переменных была оставлена в модели, если исключение ее из модели ухудшало качество (по информационным критериям Акаике и Шварца) и прогнозную силу модели.

Таблица 1. Описание переменных

	Россия	Германия
	Отраслевая структура (доля занятых в отрасли)	
	Сельское хозяйство, охота и лесное хозяйство; рыболовство, рыбоводство	Сельское хозяйство, лесное хозяйство и рыболовство
	Добыча полезных ископаемых	
	Обрабатывающие производства	Производственная промышленность (включая обрабатывающую промышленность без строительства)
	Производство и распределение электроэнергии, газа и воды	Обрабатывающая промышленность
	Строительство	Строительство
	Оптовая и розничная торговля; ремонт автотранспортных средств, мотоциклов, бытовых изделий и предметов личного пользования	Торговля, гостиницы и рестораны, транспорт
	Гостиницы и рестораны	Финансы, аренда, бизнес-услуги
	Транспорт и связь	
	Операции с недвижимым имуществом, аренда и предоставление услуг	
	Образование	

Окончание табл. 1

	Россия	Германия
	Здравоохранение и предоставление социальных услуг	
	Предоставление прочих коммунальных, социальных и персональных услуг	Предоставление государственных и частных услуг
	Другие виды деятельности	
	Состав занятого населения по возрастным группам	
С 15 до 19 лет	Доля занятых в возрасте 15–19 лет	—
С 20 до 29 лет	Доля занятых в возрасте 20–29 лет	Доля людей в возрасте от 15 до 25 лет в населении трудоспособного возраста (15–65)
С 60 до 72 лет	Доля занятых в возрасте 60–72 лет	Доля людей в возрасте от 50 до 65 лет в населении трудоспособного возраста (15–65)
	Уровень образования	
Высшее	Доля занятых с высшим образованием	Доля занятых с университетским образованием в общем числе занятых
Среднее профессиональное	—	Доля занятых с профессиональным образованием в общем числе занятых
Только школьное	—	Доля занятых без профессионального образования в общем числе занятых (только школьное образование)
	Другие факторы	
Городское население	Удельный вес городского населения в общей численности населения	—
Миграция	Коэффициент миграционного прироста на 10 тыс. человек населения	Коэффициент миграции
Доступность жилья	Средние цены на вторичном рынке жилья (руб. за кв. метр), скорректированные на среднедушевые денежные доходы населения (руб. в месяц)	—
Рост занятости	—	Темп роста занятого населения в регионе
Преступность	Число зарегистрированных преступлений на 100 тыс. чел. населения	—
Валовый региональный продукт	ВРП на душу населения в ценах 2005 года, скорректированный на стоимость потребительской корзины	ВРП на душу населения региона, деленный на индекс потребительских цен федеральной земли, в которой находится регион
Плотность	Плотность населения	Плотность населения
Заработная плата	Среднемесячная номинальная начисленная заработная плата работников организаций (в рублях)	Средняя почасовая заработная плата в регионе

3.3. Оценка моделей

При прогнозировании по моделям с объясняющими переменными необходимо знать значения этих переменных в точках прогноза. Поэтому использовались две следующие возможности:

1) оценивались модели 1–11 (см. раздел 3.1) по данным за 2005–2010 гг., подставлялись реальные значения X_t за 2011 и 2012 гг. (для России) и за 2011 г. (для Германии) и сравнивались полученные прогнозы;

2) модели 1–11 (из раздела 3.1) оценивались с заменой X_t на X_{t-1} по данным за 2006–2012 гг. (для России) и за 2006–2011 гг. (для Германии), с использованием значения X за 2012 г. (для России) и 2011 г. (для Германии) в формулах для прогноза на 2013 г. (для России) и 2012 г. (для Германии).

Пространственные модели панельных данных можно оценивать несколькими способами: с помощью обобщенного метода моментов, метода максимального правдоподобия, метода Монте-Карло на марковских цепях (МСМС). Все эти методы (при выполнении определенных условий) позволяют получить состоятельные оценки. Однако применение метода моментов затрудняется необходимостью поиска валидных инструментов. Метод Монте-Карло является достаточно трудо- (и время-) затратным. Поэтому использовался метод максимального правдоподобия, в настоящее время широко применяемый при оценке пространственных моделей панельных данных (Anselin, 1988). В ходе оценки использовалась коррекция Lee, Yu (2010), учитывающая, что T не стремится к бесконечности. Модели, не учитывающие пространственные эффекты, оценивались стандартными методами оценки панельных данных. Модель с фиксированными эффектами — с помощью оценки «within», модель со случайными эффектами — с помощью взвешенного МНК, объединенная модель регрессии — с помощью МНК.

В целях учета пространственных взаимосвязей между регионами использовались поочередно две экзогенные пространственные матрицы W: граничная матрица и матрица обратных расстояний. В граничной матрице элемент, соответствующий двум регионам, равен 1, если регионы имеют общую границу, и 0 в противном случае. Матрица обратных расстояний состоит из элементов, равных обратному расстоянию между столицами регионов (для России — по автодорогам, для Германии — прямое расстояние). Диагональные элементы матриц равны нулю. Матрицы нормированы по строкам.

В настоящей статье в целях экономии места опускаем результаты оценки моделей[4].

4. Прогнозирование

4.1. Формулы для прогнозов

В данном разделе рассматриваются формулы для построения всех представленных в статье прогнозов, начиная с наиболее простых и переходя к более сложным.

В качестве одного из наивных прогнозов использовался уровень безработицы в предыдущем периоде:

[4] Результаты доступны по запросу: lena.sem7@mail.ru.

$$\hat{y}_{i,T+s} = y_{i,T}, \quad s \geq 1. \tag{2}$$

Другой наивный прогноз вычислялся как среднее значений уровня безработицы четырех предыдущих периодов[5]:

$$\hat{y}_{i,T+s} = \sum_{j=1}^{4} y_{i,T+s-j} / 4. \tag{3}$$

Здесь, если $s > j$, то $y_{i,T+s-j} = \hat{y}_{i,T+s-j}$.

Прогноз пространственной модели с фиксированными эффектами, не включающей объясняющие переменные:

$$\hat{y}_{i,T+s} = (I - \hat{\rho} W)^{-1} \hat{\mu}_i + (I - \hat{\rho} W)^{-1} \phi_{T+s}. \tag{4}$$

Прогноз динамической пространственной модели, не включающей объясняющие переменные:

$$\hat{y}_{i,T+s} = (I - \hat{\rho} W)^{-1} \hat{\mu}_i + (I - \hat{\rho} W)^{-1} \hat{\gamma} y_{i,T+s-1} + (I - \hat{\rho} W)^{-1} \hat{\phi}_{T+s}. \tag{5}$$

Далее рассмотрим прогнозы, построенные с помощью моделей панельных данных: моделей с фиксированными и случайными эффектами с учетом и без учета пространственного лага. Для их построения необходимо располагать матрицей регрессоров для прогнозируемых периодов $X_{i,T+s}$. Предсказание уровня безработицы в регионе с помощью модели с фиксированными эффектами вычислялось следующим образом:

$$\hat{y}_{i,T+s} = X_{i,T+s} \hat{\beta} + \hat{\mu}_i + \hat{\phi}_{T+s}, \tag{6}$$

где $\hat{\mu}_i$ и $\hat{\beta}$ — within-оценки соответствующих параметров, $\hat{\phi}_{T+s}$ — оценка временного эффекта для периода $T + s$.

Из результатов оценки моделей известны оценки временных эффектов лишь для периода оценивания, поэтому в качестве $\hat{\phi}_{T+s}$ добавлялось либо значение временного эффекта предыдущего периода, либо среднее значение оценок временных эффектов нескольких (в нашем случае четырех) предыдущих периодов.

Прогноз для модели с фиксированными эффектами и пространственным авторегрессионным лагом (SAR) вычислялся с помощью формулы

$$\hat{y}_{i,T+s} = (I - \hat{\rho} W)^{-1} X'_{i,T+s} \hat{\beta} + (I - \hat{\rho} W)^{-1} \hat{\mu}_i + (I - \hat{\rho} W)^{-1} \hat{\phi}_{T+s}, \tag{7}$$

где $\hat{\rho}$, $\hat{\beta}$, $\hat{\mu}_i$ — оценки, полученные методом максимального правдоподобия для пространственных моделей панельных данных, $\hat{\phi}_{T+s}$ — оценка временного эффекта для периода $T + s$.

Для построения прогноза в моделях со случайными эффектами использовалась формула для наилучшего предсказания в классе линейных несмещенных оценок (BLUP) для ОМНК модели (Goldberger, 1962). На основе этой формулы были выведены предсказания для модели с индивидуальными эффектами (Taub, 1979; Wansbeek, Kapteyn, 1978). Позже Baltagi, Li (2004) адаптировали эту формулу для модели с пространственной корреляцией в ошиб-

[5] Были рассмотрены также средние значения двух, трех, пяти и шести предыдущих периодов. Однако в таблицу для сравнения моделей вынесен только один показатель, дающий наилучшее качество прогноза.

ках. В настоящей статье использовалась следующая формула для модели со случайными эффектами и пространственным авторегрессионным лагом:

$$\hat{y}_{i,T+s} = (I - \hat{\rho} W)^{-1} X'_{i,T+s} \hat{\beta} + (I - \hat{\rho} W)^{-1} \hat{\phi}_{T+s} + (I - \hat{\rho} W)^{-1} \frac{T\sigma_\mu^2}{\sigma_1^2} \overline{\varepsilon}_{i,OMHK}, \quad (8)$$

где $\hat{\sigma}_1^{\,2} = T\hat{\sigma}_\mu^2 + \hat{\sigma}_\varepsilon^2$, $\overline{\varepsilon}_{i,OMHK} = \sum_{t=1}^{T} \hat{\varepsilon}_{it,OMHK} / T$.

4.2. Меры для сравнения качества прогнозов

Предсказательная сила разных моделей для всех регионов в совокупности сравнивалась с помощью нескольких критериев, основным из которых является средняя квадратичная ошибка (RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} e_i^2}, \quad (9)$$

где e_i — ошибка прогноза, вычисляемая как разность между предсказанными и наблюдаемыми значениями $\hat{y}_i$ и y_i. Наряду с RMSE, наиболее часто используемым в литературе, вычислены и другие меры ошибки предсказания:

— средняя абсолютная ошибка в процентах (MAPE)

$$MAPE = 100 \frac{1}{N} \sum_{i=1}^{N} \left| \hat{y}_{i,T+s} - y_{iT+s} \right| \Big/ \left| y_{i,T+s} \right|, \quad (10)$$

— средняя абсолютная ошибка (MAE)

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| \hat{y}_{i,T+s} - y_{i,T+s} \right|, \quad (11)$$

— симметричная средняя абсолютная ошибка в процентах (sMAPE)

$$sMAPE = 100 \frac{1}{N} \sum_{i=1}^{N} 2 \left| \hat{y}_{i,T+s} - y_{i,T+s} \right| \Big/ \left| \hat{y}_{i,T+s} + y_{i,T+s} \right|. \quad (12)$$

Для оценки качества подгонки моделей для каждого региона по отдельности использовался показатель $MAE_i = \left| \hat{y}_{i,T+s} - y_{i,T+s} \right|$.

4.3. Результаты прогнозирования

Как описано выше, строились прогнозы для двух случаев: когда объясняющие переменные в будущем периоде известны, и можно было строить прогнозы в соответствующих точках, и когда объясняющие переменные в будущем периоде неизвестны, тогда приходилось использовать их лаги.

1. В первом случае оценивались модели 1–11, в предположении доступности значений соответствующих регрессоров. Результаты оценки качества прогнозов для российских регионов на один год вперед представлены в табл. 2. 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: с $RMSE = 1.48$ 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. табл. 4) 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 наилучшим прогнозом для западных регионов обладает пространственная динамическая модель с фиксированными эффектами без объясняющих переменных, а для восточных регионов — пространственная динамическая модель с фиксированными эффектами.

Результаты оценки качества предсказания на два шага вперед для российских регионов[6] (см. табл. 3) 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.

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

В таблицах 4 и 5 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 непространственными моделями панельных данных. Однако наилучшим предсказанием в данном случае

[6] Прогноз на два шага вперед для регионов Германии не был построен в связи с недостатком данных для некоторых регрессоров.

Таблица 2. Меры качества прогноза на 2011 г. для российских регионов

	Модели без регрессоров				Непространственные модели с регрессорами					Пространственные модели											
	Н	НСр	ПФЭ	ДФЭ	Скв	СЭ	ФЭ	СЭ	ФЭ	СЭ	ФЭ	СЭ	ФЭ	ДФЭ	ДФЭ	СЭ	ФЭ	СЭ	ФЭ	ДФЭ	ДФЭ
Временные эффекты	Нет	Нет	Нет	Нет	Нет	Нет	Нет	Да	Да	Нет	Нет	Да	Да	Нет	Да	Нет	Нет	Да	Да	Нет	Да
Матрица	—	—	ОР	—	—	—	—	—	—	ОР	ОР	ОР	ОР	ОР	ОР	Гр	Гр	Гр	Гр	Гр	Гр
Все регионы																					
RMSE	2.19	1.32	1.16	1.30	2.23	2.01	1.48	1.99	2.26	1.84	**0.99**	2.70	2.50	1.03	2.64	1.80	1.05	2.23	1.92	1.24	2.51
MAPE	29.13	14.11	11.23	13.74	29.83	28.68	15.23	23.26	28.49	21.07	**9.79**	36.82	36.51	11.19	38.46	20.34	10.57	28.08	26.35	12.86	36.24
MAE	1.77	0.97	0.82	0.96	1.82	1.64	1.08	1.52	1.86	1.39	**0.71**	2.34	2.30	0.80	2.44	1.37	0.76	1.81	1.67	0.92	2.31
sMAPE	23.83	13.03	10.66	12.59	24.58	22.95	13.83	19.83	23.77	18.18	**9.50**	29.52	29.51	11.56	30.98	17.74	10.00	23.36	22.25	11.90	29.46
Западные регионы																					
RMSE	2.11	1.34	1.22	1.28	2.02	7.07	1.32	2.61	3.64	1.61	**1.00**	3.27	3.40	3.22	1.50	1.69	**0.98**	2.30	2.34	3.10	1.54
MAPE	28.77	14.45	11.64	13.89	28.23	92.04	15.25	35.72	58.22	17.42	**12.01**	49.26	56.12	52.62	21.43	18.72	**11.23**	30.46	36.81	50.36	22.05
MAE	1.72	0.96	0.82	0.92	1.64	5.75	1.03	2.20	3.43	1.14	**0.78**	2.97	3.26	3.08	1.29	1.23	**0.75**	1.90	2.14	2.95	1.33
sMAPE	24.03	13.34	11.08	12.72	27.29	72.09	14.41	28.86	42.88	15.57	**12.75**	37.75	41.61	39.75	18.71	16.59	**11.02**	25.23	29.45	38.38	19.14
Восточные регионы																					
RMSE	1.26	1.26	0.68	1.37	1.26	2.56	0.99	2.04	4.64	1.23	**0.54**	0.94	1.66	1.07	1.62	1.11	0.98	0.94	1.83	0.96	1.85
MAPE	12.83	13.35	4.56	13.99	13.08	31.29	9.94	23.53	59.15	13.24	**3.60**	9.51	19.51	10.83	17.92	11.92	5.11	8.75	21.98	8.89	21.32
MAE	0.96	1.01	0.35	1.08	0.95	2.28	0.75	1.80	4.52	1.01	**0.30**	0.78	1.48	0.85	1.39	0.93	0.46	0.73	1.66	0.72	1.63
sMAPE	12.03	12.33	4.30	12.73	12.28	26.54	9.27	27.79	89.02	12.11	**3.72**	9.55	22.28	10.05	20.46	11.12	5.06	8.53	25.46	8.48	24.84

Примечание. Корректировка прогноза в связи с наличием временных эффектов модели вычисляется как среднее четырех предыдущих значений временных эффектов.

Н — наивный прогноз, НСр — наивный прогноз (по среднему), СЭ — случайные эффекты, ФЭ — фиксированные эффекты, ПФЭ — пространственная модель с фиксированными эффектами, ДФЭ — динамическая модель с фиксированными эффектами, Скв — модель сквозной регрессии, ОР — матрица обратных расстояний, Гр — граничная матрица. Жирным шрифтом выделены прогнозы наилучшего качества.

Таблица 3. Меры качества прогноза на 2012 г. для российских регионов

	Модели без регрессоров				Непространственные модели с регрессорами					Пространственные модели										
	Н	НСр	ПФЭ	ДФЭ	Скв	СЭ	ФЭ	СЭ	ФЭ	СЭ	ФЭ	СЭ	ФЭ	ДФЭ	ДФЭ	СЭ	ФЭ	СЭ	ФЭ	ДФЭ	ДФЭ
Временные эффекты	Нет	Нет	Нет	Нет	Нет	Нет	Нет	Да	Да	Нет	Нет	Да	Да	Нет	Да	Нет	Нет	Да	Да	Нет	Да
Матрица	—	—	ОР	—	—	—	—	—	—	ОР	ОР	ОР	ОР	ОР	ОР	Гр	Гр	Гр	Гр	Гр	Гр
Все регионы																					
RMSE	2.97	2.34	1.80	2.07	2.42	2.44	2.43	2.57	3.43	2.12	**1.48**	3.14	3.71	1.76	4.24	2.29	1.69	2.77	3.04	1.95	4.03
MAPE	51.04	40.89	26.69	32.16	42.26	38.82	36.32	41.86	57.28	32.61	**21.38**	55.73	68.71	29.86	78.62	36.05	24.40	47.34	54.12	27.34	74.03
MAE	2.55	2.09	1.42	1.67	2.02	1.98	1.99	2.12	3.07	1.68	**1.17**	2.79	3.53	1.48	4.02	1.85	1.35	2.38	2.80	1.53	3.79
sMAPE	36.95	31.78	22.21	25.80	31.39	30.11	30.70	31.93	42.00	25.96	**19.38**	40.03	48.10	27.81	52.68	28.24	20.90	35.14	40.16	22.92	50.44
Западные регионы																					
RMSE	2.76	2.37	1.60	2.18	2.06	2.07	1.80	3.15	4.89	1.58	**0.91**	1.83	3.91	3.43	5.14	2.01	**1.14**	1.84	2.96	1.58	4.92
MAPE	51.87	43.84	20.64	36.18	38.16	35.18	29.31	58.92	99.09	21.57	**13.33**	26.55	67.72	73.59	106.93	31.46	**15.24**	27.02	49.43	24.96	101.85
MAE	2.51	2.16	1.04	1.75	1.80	1.66	1.45	2.81	4.69	1.10	**0.58**	1.34	3.16	3.17	4.96	1.51	**0.75**	1.36	2.34	1.23	4.74
sMAPE	43.52	34.31	16.97	28.54	30.24	27.24	32.05	42.25	62.37	18.59	**32.33**	21.91	42.68	74.40	65.05	24.79	**16.07**	22.21	34.32	26.79	63.01
Восточные регионы																					
RMSE	1.75	1.86	0.83	1.86	1.61	1.53	1.58	1.71	3.98	1.57	**0.64**	1.02	1.13	1.54	1.48	1.53	**0.64**	1.07	1.33	1.45	1.83
MAPE	22.76	25.33	6.05	24.57	19.99	19.44	17.81	21.24	56.36	19.98	**4.02**	10.66	14.48	17.30	19.01	18.89	**3.92**	10.97	17.48	15.26	24.28
MAE	1.43	1.60	0.38	1.55	1.25	1.25	1.16	1.45	3.84	1.30	**0.26**	0.76	0.97	1.14	1.27	1.24	**0.26**	0.78	1.17	1.04	1.61
sMAPE	19.56	21.32	5.25	20.71	17.38	16.87	15.37	25.21	86.24	17.37	**3.60**	10.43	16.14	14.99	21.99	16.51	**3.53**	10.48	19.94	13.47	29.25

Примечание. Корректировка прогноза в связи с наличием временных эффектов модели вычисляется как среднее четырех предыдущих значений временных эффектов.

Н — наивный прогноз, НСр — наивный прогноз (по среднему), СЭ — случайные эффекты, ФЭ — фиксированные эффекты, ПФЭ — пространственная модель с фиксированными эффектами, ДФЭ — динамическая модель с фиксированными эффектами, Скв — модель сквозной регрессии, ОР — матрица обратных расстояний, Гр — граничная матрица. Жирным шрифтом выделены прогнозы наилучшего качества.

Таблица 4. Меры качества прогноза на 2011 г. для немецких регионов

	Модели без регрессоров				Непространственные модели с регрессорами					Пространственные модели										
	Н	НСр	ПФЭ	ДФЭ	Скв	СЭ	ФЭ	СЭ	ФЭ	СЭ	ФЭ	СЭ	ФЭ	ДФЭ	ДФЭ	СЭ	ФЭ	СЭ	ФЭ	ДФЭ	ДФЭ
Временные эффекты	Нет	Нет	Нет	Нет	Нет	Нет	Нет	Да	Да	Нет	Нет	Да	Да	Нет	Да	Нет	Нет	Да	Да	Нет	Да
Матрица	—	—	ОР	—	—	—	—	—	—	ОР	ОР	ОР	ОР	ОР	ОР	Гр	Гр	Гр	Гр	Гр	Гр
Все регионы																					
RMSE	0.76	1.28	1.96	0.58	2.57	1.32	1.54	1.00	0.89	5.15	2.27	13.36	3.42	**0.45**	3.78	1.46	1.28	2.11	1.12	0.58	0.78
MAPE	13.97	19.97	32.46	7.64	44.38	23.20	26.23	14.60	12.28	105.28	46.17	268.43	66.80	**6.15**	75.64	26.93	23.29	38.58	17.85	7.97	14.32
MAE	0.67	1.09	1.78	0.42	2.09	1.10	1.22	0.79	0.68	5.11	2.21	13.35	3.38	**0.33**	3.76	1.24	1.08	2.00	0.97	0.42	0.70
sMAPE	12.70	17.74	27.32	7.81	33.78	42.75	21.07	13.14	11.33	164.75	63.91	109.00	48.33	**6.10**	173.08	28.32	36.03	31.31	15.90	8.37	13.02
Западные регионы																					
RMSE	0.73	1.00	1.63	**0.44**	1.88	1.13	1.24	1.59	1.55	2.25	1.09	3.47	2.34	**0.46**	5.43	1.15	0.94	1.57	1.80	0.49	0.53
MAPE	14.63	19.21	31.96	**6.69**	31.41	20.08	20.26	30.76	29.71	48.08	22.23	72.69	48.44	**7.22**	116.95	22.04	17.04	30.07	36.10	7.10	9.84
MAE	0.66	0.91	1.55	**0.33**	1.45	0.93	0.93	1.47	1.43	2.17	0.99	3.42	2.28	**0.35**	5.42	0.98	0.76	1.44	1.71	0.34	0.45
sMAPE	13.27	17.12	26.93	**6.53**	28.32	24.65	38.31	25.78	25.07	28.97	27.10	51.47	37.86	**6.96**	388.74	28.04	20.95	25.30	29.74	7.35	9.19
Восточные регионы																					
RMSE	0.93	2.50	3.48	0.80	1.94	1.07	1.14	2.21	2.53	1.21	1.20	1.78	2.46	1.17	1.14	1.08	0.97	2.29	1.51	**0.78**	2.03
MAPE	9.29	25.34	35.98	6.57	15.94	9.24	9.09	21.53	25.06	10.54	10.51	16.58	24.72	10.68	10.72	9.29	7.69	22.40	14.05	**6.41**	20.84
MAE	0.80	2.39	3.40	0.64	1.47	0.85	0.85	2.05	2.39	0.99	0.99	1.60	2.36	1.02	1.00	0.85	0.72	2.14	1.35	**0.63**	1.94
sMAPE	8.63	22.10	30.03	6.87	15.05	9.95	9.78	19.09	21.91	11.52	11.43	15.03	21.71	11.49	9.99	10.05	8.17	19.80	12.91	**6.57**	18.55

Примечание. Корректировка прогноза в связи с наличием временных эффектов модели вычисляется как среднее четырех предыдущих значений временных эффектов.

Н — наивный прогноз, НСр — наивный прогноз (по среднему), СЭ — случайные эффекты, ФЭ — фиксированные эффекты, ПФЭ — пространственная модель с фиксированными эффектами, ДФЭ — динамическая модель с фиксированными эффектами, Скв — модель сквозной регрессии, ОР — матрица обратных расстояний, Гр — граничная матрица. Жирным шрифтом выделены прогнозы наилучшего качества.

Таблица 5. Меры качества прогноза на 2013 г. для российских регионов

	Модели без регрессоров				Непространственные модели с регрессорами					Пространственные модели										
	Н	НСр	ПФЭ	ДФЭ	Скв	СЭ	ФЭ	СЭ	ФЭ	СЭ	ФЭ	СЭ	ФЭ	ДФЭ	ДФЭ	СЭ	ФЭ	СЭ	ФЭ	ДФЭ	ДФЭ
Временные эффекты	Нет	Нет	Нет	Нет	Нет	Нет	Нет	Да	Да	Нет	Нет	Да	Да	Нет	Да	Нет	Нет	Да	Да	Нет	Да
Матрица	—	—	ОР	—	—	—	—	—	—	ОР	ОР	ОР	ОР	ОР	ОР	Гр	Гр	Гр	Гр	Гр	Гр
Все регионы																					
RMSE	**0.72**	1.77	1.83	1.05	1.53	1.14	1.25	1.27	1.62	1.00	1.15	1.13	1.88	0.77	**0.73**	1.10	1.20	1.09	1.67	1.03	0.75
MAPE	**10.09**	28.94	27.02	17.22	23.08	16.42	18.36	20.88	32.50	14.49	16.94	16.91	32.28	11.37	**11.09**	15.14	17.43	16.03	28.51	15.19	11.17
MAE	**0.55**	1.54	1.48	0.89	1.23	0.90	0.94	1.19	1.66	0.80	0.88	0.93	1.66	0.59	**0.57**	0.85	0.91	0.88	1.46	0.83	0.57
sMAPE	**10.21**	24.12	22.57	21.77	21.19	15.29	20.56	21.62	88.10	15.32	18.39	19.43	22.68	11.33	**11.58**	14.16	17.82	17.81	9.97	17.73	11.36
Западные регионы																					
RMSE	**0.68**	1.84	1.92	1.09	1.37	1.04	1.26	1.05	1.52	1.01	1.62	1.10	2.40	**0.86**	0.99	1.06	1.31	1.14	1.84	1.28	0.86
MAPE	**10.54**	31.51	29.03	20.48	22.24	15.71	20.76	16.56	28.06	16.02	30.27	18.32	45.52	**13.61**	17.98	16.24	22.29	18.53	34.30	20.41	15.20
MAE	**0.53**	1.59	1.54	0.95	1.13	0.82	0.95	0.84	1.30	0.80	1.41	0.88	2.15	**0.65**	0.86	0.83	1.05	0.91	1.60	1.05	0.73
sMAPE	**10.84**	25.99	24.20	27.09	25.16	15.70	29.63	17.63	65.95	16.32	30.58	19.94	46.38	**12.76**	22.31	15.87	30.81	19.12	8.20	24.77	17.81
Восточные регионы																					
RMSE	**0.81**	1.60	1.61	4.48	1.44	1.28	1.50	0.98	0.98	0.90	1.14	**0.84**	0.86	1.12	0.95	0.93	1.19	0.85	0.88	1.12	1.10
MAPE	**9.06**	23.12	22.46	65.89	19.54	21.32	27.13	12.76	12.73	10.96	15.16	**10.73**	9.32	14.51	10.33	11.02	15.71	10.83	9.37	14.51	11.86
MAE	**0.60**	1.42	1.32	4.41	1.18	1.18	1.39	0.76	0.77	0.70	0.94	**0.65**	0.65	0.91	0.76	0.72	0.98	0.67	0.66	0.91	0.89
sMAPE	**8.80**	19.89	18.89	104.37	17.02	19.82	32.10	11.41	11.85	10.12	13.73	**9.86**	9.18	13.23	10.90	10.16	14.14	10.00	9.33	13.23	12.87

Примечание. Корректировка прогноза в связи с наличием временных эффектов модели вычисляется как среднее четырех предыдущих значений временных эффектов.

Н — наивный прогноз, НСр — наивный прогноз (по среднему), СЭ — случайные эффекты, ФЭ — фиксированные эффекты, ПФЭ — пространственная модель с фиксированными эффектами, ДФЭ — динамическая модель с фиксированными эффектами, Скв — модель сквозной регрессии, ОР — матрица обратных расстояний, Гр — граничная матрица. Жирным шрифтом выделены прогнозы наилучшего качества.

Таблица 6. Меры качества прогноза на 2012 г. для немецких регионов

	Модели без регрессоров				Непространственные модели с регрессорами					Пространственные модели											
	Н	НСр	ПФЭ	ДФЭ	Скв	СЭ	ФЭ	СЭ	ФЭ	СЭ	ФЭ	СЭ	ФЭ	ДФЭ	ДФЭ	СЭ	ФЭ	СЭ	ФЭ	ДФЭ	ДФЭ
Временные эффекты	Нет	Нет	Нет	Нет	Нет	Нет	Нет	Да	Да	Нет	Нет	Да	Да	Нет	Да	Нет	Нет	Да	Да	Нет	Да
Матрица	—	—	ОР	—	—	—	—	—	—	ОР	ОР	ОР	ОР	ОР	ОР	Гр	Гр	Гр	Гр	Гр	Гр
Все регионы																					
RMSE	**0.34**	0.95	2.16	0.74	2.27	0.70	1.33	2.88	2.64	4.15	1.22	10.35	1.67	0.49	5.90	0.80	0.84	2.86	3.13	0.52	**0.39**
MAPE	**4.25**	15.11	36.35	14.20	40.43	10.09	24.29	59.67	54.60	87.50	24.73	217.13	34.63	7.83	122.65	14.40	14.15	60.08	65.88	8.22	**5.76**
MAE	**0.24**	0.80	1.94	0.67	1.84	0.53	1.09	2.82	2.57	4.12	1.11	10.33	1.57	0.39	5.89	0.66	0.67	2.80	3.06	0.41	**0.31**
sMAPE	**4.13**	13.75	30.00	15.87	31.45	10.44	51.71	107.53	87.26	211.41	32.69	275.41	2.40	8.25	885.70	16.98	16.98	108.33	139.6	8.74	**5.95**
Западные регионы																					
RMSE	**0.31**	0.76	1.79	1.05	1.69	0.95	1.03	2.35	2.15	1.39	0.72	0.55	0.58	0.53	4.95	0.62	0.71	2.26	2.73	0.46	**0.40**
MAPE	**4.22**	14.60	35.64	20.67	28.79	17.84	18.86	51.76	47.37	30.06	13.89	9.84	9.87	9.86	110.23	11.23	12.21	49.83	60.71	8.12	**6.85**
MAE	**0.22**	0.67	1.68	0.95	1.30	0.80	0.82	2.30	2.09	1.31	0.61	0.45	0.46	0.44	4.94	0.51	0.56	2.20	2.67	0.37	**0.32**
sMAPE	**4.09**	13.32	29.48	18.28	25.57	21.08	23.88	73.10	160.03	39.04	15.66	10.49	10.30	10.63	344.71	12.21	13.46	66.68	105.2	8.58	**7.18**
Восточные регионы																					
RMSE	**0.52**	1.78	1.46	0.61	1.67	1.15	1.12	3.62	4.24	0.80	0.78	2.65	3.18	0.98	0.91	0.94	0.96	3.43	3.82	0.84	**0.80**
MAPE	**4.52**	18.66	5.86	5.52	14.06	11.00	10.98	39.21	46.32	6.90	6.28	28.18	34.45	9.50	8.73	8.11	8.05	37.15	41.70	7.96	**5.91**
MAE	**0.44**	1.68	0.54	0.51	1.28	0.94	0.94	3.54	4.17	0.64	0.60	2.54	3.08	0.84	0.77	0.75	0.74	3.34	3.73	0.69	**0.57**
sMAPE	**4.39**	16.76	4.78	5.69	13.84	12.20	12.07	51.01	63.74	7.13	6.25	34.03	43.51	10.18	9.27	7.90	7.97	47.77	55.80	8.36	**5.65**

Примечание. Корректировка прогноза в связи с наличием временных эффектов модели вычисляется как среднее четырех предыдущих значений временных эффектов.

Н — наивный прогноз, НСр — наивный прогноз (по среднему), СЭ — случайные эффекты, ФЭ — фиксированные эффекты, ПФЭ — пространственная модель с фиксированными эффектами, ДФЭ — динамическая модель с фиксированными эффектами, Скв — модель сквозной регрессии, ОР — матрица обратных расстояний, Гр — граничная матрица. Жирным шрифтом выделены прогнозы наилучшего качества.

является наивное предсказание: 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 на 2012, 2011 и 2010 гг. (для России) и 2011, 2010 гг. (для Германии). Для 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 Германии). В Приложении 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).

Таким образом, включение пространственных эффектов позволяет получить лучшее качество предсказания безработицы в региональном разрезе в сравнении с непространственными моделями панельных данных. Настоящее исследование дополнительно отвечает на вопрос, имеет ли смысл использовать сложные пространственные модели панельных данных вместо простых в построении наивных предсказаний, часто используемых на 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.

5. Заключение

В работе осуществлена попытка прогнозирования уровня безработицы в регионах России и Германии на основе пространственных моделей панельных данных. Цель исследования — показать, как учет пространственных эффектов в модели влияет на качество прогноза, а также ответить на вопрос, позволяют ли пространственные эконометрические модели получить лучшее качество прогноза, чем наивные прогнозы.

Качество прогнозов, полученных с помощью различных моделей, проверялось для двух случаев. В первом случае в качестве значений регрессоров для прогнозов использовались их известные значения, а во втором — их динамические лаги (в связи с отсутствием таких значений). Результаты оценки качества прогнозов показали, что включение пространственных эффектов в модель региональной безработицы значительно улучшает предсказание по сравнению с непространственными моделями панельных данных. Кроме того, при известных значениях объясняющих переменных прогнозы, полученные с помощью пространственных моделей, дают лучшее качество предсказания, в том числе и по сравнению с наивными прогнозами, часто используемыми в практике. Даже в случае отсутствия прогнозных будущих значений регрессоров и использованием их динамических лагов, пространственные

модели лишь незначительно уступают наивным предсказаниям, а для некоторых отдельных регионов позволяют получить даже лучшее качество прогноза.

Благодарности. Авторы выражают благодарность Р. Косфельду, С. Гондве и другим участникам секции «Рынки труда» 10-й международной конференции ассоциации пространственной эконометрики (15 июня 2016 г., Рим), Г. Канторовичу, Ю. Ачкасову, Т. Ратниковой, Д. Малахову, Н. Пильнику и другим участникам семинара лаборатории макроструктурного моделирования экономики России (12 января 2016 г., Москва), И. Антохоновой, А. Березняцкому, А. Шемякину и другим участникам IX-й Международной школы-семинара «Многомерный статистический анализ и эконометрика» (30 июня 2016 г., Республика Армения), а также анонимному рецензенту за сделанные критические замечания и высказанные предложения, многие из которых авторы учли в финальной версии статьи, а оставшиеся постараются реализовать в дальнейших исследованиях.

Список литературы

Ачкасов Ю., Шатило Е. (2016). Рынок труда: Положение, прогноз, развитие. *Научный вестник ИЭП им. Гайдара.ру*, 1 (95), 53–55.

Гурвич Е., Вакуленко Е. (2016). Подходы к моделированию российского рынка труда. *Научный вестник ИЭП им. Гайдара.ру*, 1 (95), 40–43.

Регионы России. Социально-экономические показатели. Стат. сборники. М.: Росстат. 2005–2013.

Турунцева М., Астафьева Е. (2016). Прогнозирование занятости и безработицы в краткосрочном периоде: Некоторые подходы. *Научный вестник ИЭП им. Гайдара.ру*, 1 (95), 58–60.

Anselin L. (1988). *Spatial econometrics: Methods and models*. Kluwer Academic Publishers, Dordrecht.

Aragon Y., Haughton D., Haughton J., Leconte E., Malin E., Ruiz-Gazen A., Thomas-Agnan C. (2003). Explaining the pattern of regional unemployment: The case of the Midi-Pyrénées region. *Papers in Regional Science*, 82, 155–174.

Armstrong H., Taylor J. (1988). Regional policy and the North-South divide. In: *Making the Economy Work*, Palgrave Macmillan UK, 125–172.

Baltagi B. H., Fingleton B., Pirotte A. (2014). Estimating and forecasting with a dynamic spatial panel data model. *Oxford Bulletin of Economics and Statistics*, 76, 112–138.

Baltagi B. H., Li D. (2004). Prediction in the panel data models with spatial correlation. In: *Advances in spatial econometrics*, Springer, 283–295.

Baltagi B. H., Li D. (2006). Prediction in the panel data model with spatial correlation: The case of liquor. *Spatial Economic Analysis*, 1, 175–185.

Elhorst J. P. (2003). The mystery of regional unemployment differentials: Theoretical and empirical explanations. *Journal of economic surveys*, 17, 709–748.

Girardin E., Kholodilin K. A. (2011). How helpful are spatial effects in forecasting the growth of Chinese provinces? *Journal of Forecasting*, 30, 622–643.

Goldberger A. S. (1962). Best linear unbiased prediction in the generalized linear regression models. *Journal of the American Statistical Association*, 57, 369–375.

Greenwood M. J., Hunt G. L., Rickman D. S., Treyz G. I. (1991). Migration, regional equilibrium, and the estimation of compensation differentials. *The American Economic Review*, 81, 1382–1390.

Hofler R. A., Murphy K. J. (1989). Using a composed error model to estimate the frictional and excess-supply components of unemployment. *Journal of Regional Science*, 29, 213–228.

Kholodilin K., Siliverstovs B., Kooths S. (2008). A dynamic panel data approach to the forecasting of the GDP of German Länder. *Spatial Economic Analysis*, 3, 195–207.

Lee L.-F., Yu J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154, 165–185.

Lehmann R., Wohlrabe K. (2014). Regional economic forecasting: State-of-the-art methodology and future challenges. *CESifo Working Paper* No. 5145.

Longhi S., Nijkamp P. (2007). Forecasting regional labor market developments under spatial autocorrelation. *International Regional Science Review*, 30, 100–119.

Lottmann F. (2012). Regional unemployment in Germany: A spatial panel data analysis. ERSA conference papers. European Regional Science Association. http://www-sre.wu.ac.at/ersa/ersaconfs/ersa12/e120821aFinal00055.pdf.

Marston S. T. (1985). Two views of the geographic distribution of unemployment. *The Quarterly Journal Economics*, 57–79.

Martin R. (1997). Regional unemployment disparities and their dynamics. *Regional Studies*, 31, 237–252.

Mayor M., Patuelli R. (2012). Short-run regional forecasts: Spatial models through varying cross-sectional and temporal dimensions. In: *Defining the Spatial Scale in Modern Regional Analysis*, Springer, 173–192.

Molho I. (1995). Spatial autocorrelation in British unemployment. *Journal of Regional Science*, 35, 641–658.

Overman H. G., Puga D., Vandenbussche H. (2002). Unemployment clusters across Europe's regions and countries. *Economic policy*, 117–147.

Partridge M. D., Rickman D. S. (1995). Differences in state unemployment rates: The role of labor and product market structural shifts. *Southern Economic Journal*, 89–106.

Polasek W., Sellner R., Schwarzbauer W. (2007). Long term regional forecasting with spatial equation systems. https://www.ihs.ac.at/publications/eco/human_resources/polasek/longtermsystempred_kurz1.pdf.

Schanne N., Wapler R., Weyh A. (2010). Regional unemployment forecasts with spatial interdependencies. *International Journal of Forecasting*, 26, 908–926.

Taub A. J. (1979). Prediction in the context of the variance-components model. *Journal of Econometrics*, 10, 103–107.

Wansbeek T. J., Kapteyn A. (1978). The separation of individual variation and systematic change in the analysis of panel data. *Annales de l'INSEE*, 659–680.

Поступила в редакцию 28.05.2016;
принята в печать 22.07.2016.

Приложение

Список регионов, для которых прогнозы на 2010–2013 гг. (для России) и 2010–2012 гг. (для Германии), полученные с помощью оценки пространственных моделей на временных промежутках 2006–2012 гг., 2006–2011 гг., 2006–2010 гг., 2006–2009 гг. (для России) и 2006–2011 гг., 2006–2010 гг., 2006–2009 гг. (для Германии), оказались наилучшими, т. е. имели наименьшее значение показателей качества прогнозов среди представленных моделей.

Регионы Германии	Регионы России
Фленсбург, город	Калужская область
Хельмштедт, район	Орловская область
Шаумбург, район	Тверская область
Дюрен, район	Ярославская область
Обербергиш, район	Республика Карелия
Брайсгау-Шварцвальд, район	Республика Коми
Ройтлинген, район	Ленинградская область
Дахау, район	Волгоградская область
Пассау, район	Пермский край
Реген, район	Алтайский край
Амберг-Зульцбах, район	Иркутская область
Хоф, район	Новосибирская область
Вунзидель-Фихтель, район	
Нойштадт-на-Айше — Бад-Виндсхайм, район	
Хафельланд, район	
Одер-Шпрее, район	
Восточный Пригниц-Руппин, район	
Шпрее-Нейсе, район	
Халле (Заале), город	
Эрфурт, город	
Веймар, город	
Айхсфельд, район	
Нордхаузен, район	
Кифхойзер, район	
Зёммерда, район	
Ильм-Крайс, район	
Сонненберг, район	
Альтенбургер, район	

Semerikova E., Demidova O. Using spatial econometric models for regional unemployment forecasting. *Applied Econometrics*, 2016, v. 43, pp. 29–51.

Elena Semerikova
National Research University Higher School of Economics, Moscow, Russian Federation;
demidova@hse.ru

Olga Demidova
National Research University Higher School of Economics, Moscow, Russian Federation;
lena.sem7@mail.ru

Using spatial econometric models for regional unemployment forecasting

We consider forecasting unemployment in Russian and German region with the help of econometric panel data models. Using regional data from 2005 till 2012 we show that spatial panel data models perform better in terms of forecasting accuracy than other models (on average and at least for some distinct regions) such as non-spatial panel data models, pooled OLS, models without exploratory variables and naive forecasts (average value for one or several previous periods).

Keywords: spatial panel data models; prediction; regional unemployment.

JEL classification: R1; C21; C23.

References

Achkasov Ju., Shatilo E. (2016). Rynok truda: Polozhenie, prognoz, razvitie. *Nauchnyj vestnik IJeP im. Gajdara.ru*, 95, 53–55 (in Russian).

Gurvich E., Vakulenko E. (2016). Podhody k modelirovaniju rossijskogo rynka truda. *Nauchnyj vestnik IJeP im. Gajdara.ru*, 95, 40–43 (in Russian).

Regiony Rossii. Socialno-ekonomicheskie pokazateli. Stat. sborniki. Rosstat. M., 2005–2013 (in Russian).

Turunceva M., Astafieva E. (2016). Prognozirovanie zanjatosti i bezraboticy v kratkosrochnom periode: Nekotorye podhody. *Nauchnyj vestnik IJeP im. Gajdara.ru*, 95, 58–60 (in Russian).

Anselin L. (1988). *Spatial econometrics: Methods and models*. Kluwer Academic Publishers, Dordrecht.

Aragon Y., Haughton D., Haughton J., Leconte E., Malin E., Ruiz-Gazen A., Thomas-Agnan C. (2003). Explaining the pattern of regional unemployment: The case of the Midi-Pyrénées region, *Papers in Regional Science*, 82, 155–174.

Armstrong H., Taylor J. (1988). Regional policy and the North-South divide. In: *Making the Economy Work*, Palgrave Macmillan UK, 125–172.

Baltagi B. H., Fingleton B., Pirotte A. (2014). Estimating and forecasting with a dynamic spatial panel data model. *Oxford Bulletin of Economics and Statistics*, 76, 112–138.

Baltagi B. H., Li D. (2004). Prediction in the panel data models with spatial correlation. In: *Advances in spatial econometrics*, Springer, 283–295.

Baltagi B. H., Li D. (2006). Prediction in the panel data model with spatial correlation: The case of liquor. *Spatial Economic Analysis*, 1, 175–185.

Elhorst J. P. (2003). The mystery of regional unemployment differentials: Theoretical and empirical explanations. *Journal of economic surveys*, 17, 709–748.

Girardin E., Kholodilin K. A. (2011). How helpful are spatial effects in forecasting the growth of Chinese provinces? *Journal of Forecasting*, 30, 622–643.

Goldberger A. S. (1962). Best linear unbiased prediction in the generalized linear regression models. *Journal of the American Statistical Association*, 57, 369–375.

Greenwood M. J., Hunt G. L., Rickman D. S., Treyz G. I. (1991). Migration, regional equilibrium, and the estimation of compensation differentials. *The American Economic Review*, 81, 1382–1390.

Hofler R. A., Murphy K. J. (1989). Using a composed error model to estimate the frictional and excess-supply components of unemployment. *Journal of Regional Science*, 29, 213–228.

Kholodilin K., Siliverstovs B., Kooths S. (2008). A dynamic panel data approach to the forecasting of the GDP of German Länder. *Spatial Economic Analysis*, 3, 195–207.

Lee L.-F., Yu J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154, 165–185.

Lehmann R., Wohlrabe K. (2014). Regional economic forecasting: State-of-the-art methodology and future challenges. *CESifo Working Paper* No. 5145.

Longhi S., Nijkamp P. (2007). Forecasting regional labor market developments under spatial autocorrelation. *International Regional Science Review*, 30, 100–119.

Lottmann F. (2012). Regional unemployment in Germany: A spatial panel data analysis. ERSA conference papers. European Regional Science Association. http://www-sre.wu.ac.at/ersa/ersaconfs/ersa12/e120821aFinal00055.pdf.

Marston S. T. (1985). Two views of the geographic distribution of unemployment. *The Quarterly Journal Economics*, 57–79.

Martin R. (1997). Regional unemployment disparities and their dynamics. *Regional Studies*, 31, 237–252.

Mayor M., Patuelli R. (2012). Short-run regional forecasts: Spatial models through varying cross-sectional and temporal dimensions. In: *Defining the Spatial Scale in Modern Regional Analysis*, Springer, 173–192.

Molho I. (1995). Spatial autocorrelation in British unemployment. *Journal of Regional Science*, 35, 641–658.

Overman H. G., Puga D., Vandenbussche H. (2002). Unemployment clusters across Europe's regions and countries. *Economic policy*, 117–147.

Partridge M. D., Rickman D. S. (1995). Differences in state unemployment rates: The role of labor and product market structural shifts. *Southern Economic Journal*, 89–106.

Polasek W., Sellner R., Schwarzbauer W. (2007). Long term regional forecasting with spatial equation systems. https://www.ihs.ac.at/publications/eco/human_resources/polasek/longtermsystempred_kurz1.pdf.

Schanne N., Wapler R., Weyh A. (2010). Regional unemployment forecasts with spatial interdependencies. *International Journal of Forecasting*, 26, 908–926.

Taub A. J. (1979). Prediction in the context of the variance-components model. *Journal of Econometrics*, 10, 103–107.

Wansbeek T. J., Kapteyn A. (1978). The separation of individual variation and systematic change in the analysis of panel data. *Annales de l'INSEE*, 659–680.

Received 28.05.2016; accepted 12.07.2016.