

ПРОГНОЗИРОВАНИЕ ИНФЛЯЦИИ РЕГИОНОВ ЕВРОПЕЙСКОЙ ЧАСТИ РОССИИ

Т.Горшкова, н.с. РАНХиГС,

М.Турунцева, зав. лабораторией ИЭП им. Е.Т. Гайдара и РАНХиГС

В статье проводится обзор эконометрических методов прогнозирования с помощью панельных данных. Целью обзора является выявление наиболее точных методов прогнозирования и ответ на вопрос о необходимости учета в модели пространственной зависимости между регионами. Также в статье приведены результаты построения модели пространственной корреляции для инфляции регионов европейской части России.

По итогам проведенного обзора сделан вывод, что результаты прогнозирования меняются в зависимости от исследуемых стран и временного периода. На российских данных нельзя дать однозначный ответ о необходимости включения в модели пространственных переменных.

1. Общие сведения о пространственных моделях

Возможное наличие пространственной корреляции является особенностью региональных / страновых данных. Модели, призванные оценить эти эффекты, требуют спецификации того, как изменение объясняемой переменной в одном регионе зависит от изменения переменных в других регионах. При моделировании данная спецификация учитывается с помощью весовой матрицы соседей W . Наиболее часто матрица весов представляет собой квадратную симметричную матрицу размера $N \times N$ с элементами (i, j) , принимающими значение 1, если регион i граничит с регионом j , и 0 – иначе. Диагональным элементам матрицы присвоено значение 0. Также существуют и другие способы построения этой матрицы (см., например, Луговой, Дашкеев и др., 2007¹). Однако до построения матрицы весов необходимо определить, присутствует ли в данных пространственная автокорреляция.

Наиболее известным *тестом на наличие пространственной автокорреляции* является I Морана, которое рассчитывается следующим образом:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (x_i - \mu)(x_j - \mu)}{\sum_i (x_i - \mu)^2},$$

где x – вектор значений объясняемой переменной, μ – их среднее, N – число наблюдений. I Морана принимает значения от -1 до 1. Значение -1 свидетельствует о том, что регионы со значениями x выше среднего соседствуют с регионами, в которых значения x ниже среднего. Значение 1 говорит о положительной корреляции: регионы, в которых x выше среднего граничат друг с другом. Если значение статистики Морана не значимо, то предполагается, что ошибки распределены между регионами случайным образом. Однако даже в этом случае учет пространственной корреляции может улучшить прогнозные свойства модели (см., например, Baltagi, B.H., Li, D., 2004², Robertson, D., Symons, J., 1992³).

¹ Луговой, Дашкеев и др. (2007). Экономико-географические и институциональные аспекты экономического роста в регионах М.: ИЭПП.

² Baltagi, B.H., Li, D., 2004. Prediction in the panel data model with spatial correlation. In: Anselin, L., Florax, R.J.G.M., Rey, S.J. (Eds.), Advances in Spatial Econometrics: Methodology, Tools and Applications. Springer, Berlin, pp. 283–295 (Chapter 13).

³ Robertson, D., Symons, J., 1992. Some strange properties of panel data estimators. Journal of Applied Econometrics 7, 175–189.

2. Обзор эмпирических работ

В работе *Baltagi, Li* (2004) рассмотрено несколько панельных моделей, с помощью которых авторы прогнозировали спрос на сигареты в 46 штатах США за период 1963–1992 гг. Для оценивания был применен ряд методов: обычный МНК, игнорирующий гетерогенность штатов и пространственную автокорреляцию. Метод максимального правдоподобия (MLE), с помощью которого оценивается модель пространственной автокорреляции. Модели с фиксированными (FE) и случайными (RE) эффектами с учетом и без учета пространственной корреляции. Наиболее точным оказался прогноз по пространственной модели с фиксированными эффектами.

Girardin, Kholodilin (2011)¹ исследуют темп роста валового регионального продукта (ВРП) в 31 провинции Китая за период 1979–2007 гг. В качестве базовой модели используется модель AR(1). С помощью пула и фиксированных эффектов (FE) оцениваются модели как учитывающие, так и неучитывающие пространственную корреляцию. Оцененные коэффициенты пространственной авторегрессии значимы и положительны, что говорит о важности пространственной зависимости среди китайских провинций. Значение R^2 в моделях, учитывающих пространственную зависимость, более чем в два раза выше по сравнению с моделями без пространственных эффектов. Все модели дают лучшие прогнозы, чем наивная модель.

В статье *Baltagi, Li* (2006)² рассматриваются цены на ликер в 43 штатах США за период 1965–1994. Цены зависят от акцизов на алкоголь, которые различаются по штатам. На основе проведенного исследования авторы делают вывод, что наилучшие прогнозы дают FE и RE оценки. Добавление в модели пространственной корреляции не улучшает качество прогнозов (за исключением прогноза на первый год).

В статье *Longhi, Nijkamp* (2007)³ сделаны краткосрочные прогнозы уровней региональной безработицы для 326 регионов Западной Германии за период с 1987–2002 гг. С помощью фиксированных и случайных эффектов оцениваются модели с учетом и без учета пространственной корреляции. По результатам моделирования можно сделать выводы, что в среднем пространственные модели дают более хорошие прогнозы, чем соответствующие непространственные модели. Единственное исключение – модель с фиксированными эффектами. Авторы, однако, делают вывод, что в целом пространственные модели не улучшают прогнозы, полученные с помощью непространственных моделей.

В работе *Marques, Pino, Horrillo* (2004)⁴ исследуется гетерогенная природа инфляционной динамики на региональном и секторальном уровнях в Чили за период 2002–2008 гг. Пространственная зависимость была обнаружена для 81% товаров; общий фактор значимо влияет только на 10% гомогенных, относительно дорогих и легко транспортируемых товаров. В связи с природными особенностями страны, транспортные издержки составляют большую часть наценки товаров, что затрудняет ценовой арбитраж между регионами.

В работе *Kholodilin, Siliverstovs, Kooths* (2008)⁵ авторы прогнозируют годовой темп роста ВВП для 16 немецких регионов за период 1992–2006 гг. В качестве базовой модели используется авторегрессия первого порядка, построенная для каждого региона отдельно. Оценки, полученные с помощью моделей, учитывающих пространственную корреляцию, практически не отличаются от оценок более простых моделей. Оценки пространственных авторегрессионных коэффициентов оказываются незначимы. Оценки модели пространственной

¹ Girardin, E., Kholodilin, K.A., 2011. How helpful are spatial effects in forecasting the growth of Chinese provinces? *Journal of Forecasting* 30, 622–643.

² 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.

³ Longhi, S., Nijkamp, P., 2007. Forecasting regional labor market developments under spatial heterogeneity and spatial correlation. *International Regional Science Review* 30, 100–119.

⁴ Marques, H., Pino, G., Horrillo J., 2014. “Regional inflation dynamics using space–time models,” *Empirical Economics*, Springer, vol. 47(3), pages 1147–1172, November.

⁵ Kholodilin, K.A., 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.

ошибки, полученные с помощью MLE, близки к МНК оценкам, но, в отличие от моделей пространственной авторегрессии, коэффициент пространственной корреляции оказывается значимым и положительным.

В работе *Arbia, Piras (2004)*¹ оценивается конвергенция темпов роста ВВП через модели панельных данных с учетом пространственной корреляции ошибки и пространственно-го лага зависимой переменной. Авторы используют данные о логарифме ВВП на душу населения, пересчитанного по паритету покупательной способности, для 125 регионов, расположенных в 10 европейских странах за 1980–1995 гг. Знак коэффициента при объясняющей переменной в пространственной модели не отвергает гипотезу о наличии конвергенции. Анализ остатков расширенной панели показывает, что пространственная корреляция не полностью описывает структуру ошибок. Коэффициент *I* Морана для остатков по-прежнему остается значимым для большинства случаев, однако его абсолютные значения снижаются, что говорит о необходимости включения пространственной корреляции в модель конвергенции.

Banerjee, Marcellino (2006) исследуют факторы, влияющие на инфляцию и ВВП в США за 1 кв. 1975 г. – 4 кв. 2001 г. При прогнозировании инфляции авторы опираются на результаты работы *Cecchetti, Chu, Steindel (2000)*², которые показали, что прогнозы, полученные с помощью авторегрессионных моделей, часто дают более точные результаты, чем структурные модели. *Banerjee, Marcellino (2006)*³ в качестве базовой модели инфляции берут авторегрессию четвертого порядка и сравнивают с ней две факторные модели: модель с финансовыми индикаторами (обменный курс, денежные агрегаты, ставка процента) и индикаторы реальной экономической активности (коэффициент использования производственных мощностей и уровень безработицы). Авторы делают вывод, что результаты, полученные с помощью базовой модели авторегрессии, в целом оказываются более точными, чем результаты факторных моделей. В связи с этим для прогнозирования предлагается использовать более простые однофакторные модели.

В работе *Yesilyurt, Elhorst (2014)*⁴ исследовались модели региональной инфляции для 67 провинций Турции за период 1987–2001 гг. Для построения моделей авторы использовали гибридную кривую Филипса, в которую включали переменные, учитывающую пространственную взаимосвязь. На основе полученных результатов авторы делают вывод, что модель с пространственной корреляцией в ошибках не учитывает влияние побочных эффектов на инфляцию.

В табл. 1 собраны результаты всех работ, рассмотренных в данном обзоре. Необходимо обратить внимание на вывод, сделанный в ряде работ, согласно которому учет пространственной корреляции не улучшает результаты, полученные с помощью непространственных моделей.

Таблица 1

РЕЗУЛЬТАТЫ РАБОТ, ПОСВЯЩЕННЫХ ПРОГНОЗИРОВАНИЮ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ

Авторы	Исследуемый показатель	Данные	Вывод
Girardin, Kholodilin (2011)	Темп роста валового регионального продукта	31 провинция Китая за период 1979–2007 гг.	Гипотеза о наличии пространственной зависимости не отвергается

¹ Arbia, G., Piras, G., 2004. “Convergence in per-capita GDP across European regions using panel data models extended to spatial autocorrelation effects,” ERSA conference papers ersa04p524, European Regional Science Association.

² Cecchetti, S., R. Chu and C. Steindel (2000): The Unreliability of Inflation Indicators, Federal Reserve Bank of New York Current Issues in Economics and Finance, 6, pp.1–6.

³ Banerjee, A., Marcellino, M., 2006. “Are there any reliable leading indicators for US inflation and GDP growth?”, International Journal of Forecasting, Elsevier, vol. 22(1), pages 137–151.

⁴ Yesilyurt, F., Elhorst, J. P. (2014). A regional analysis of inflation dynamics in Turkey. Annals of Regional Science, 52(1), 1–17. 10.1007/s00168-013-0570-4

Таблица 1, окончание

Авторы	Исследуемый показатель	Данные	Вывод
Baltagi, Li (2006)	Цены на ликер	43 штата США, 1965-1994 гг.	Добавление в модели пространственной корреляции в целом не улучшает качество прогнозов
Longhi, Nijkamp (2007)	Региональная безработица	326 регионов Западной Германии, 1987–2002 гг.	Добавление в модели пространственной корреляции в целом не улучшает качество прогнозов
Marques, Pino, Tena (2014)	Инфляционная динамика	Чили, 2002-2008 гг.	Общий фактор значимо влияет только на 10% товаров
Kholodilin, Siliverstovs, Kooths (2008)	Темп роста ВВП	16 регионов Германии, 1992–2006 гг.	Добавление в модели пространственной корреляции в целом не улучшает качество прогнозов
Arbia, Piras (2004)	Конвергенция темпов роста ВВП	10 стран Европы, 1980-1995 гг.	Пространственная корреляция не полностью описывает структуру ошибок, однако ее необходимо включать в модель
Banerjee, Marcellino (2006)	Инфляция и ВВП	США, 1 кв. 1975 – 4 кв. 2001 гг.	Базовая модель авторегрессии в целом оказывается более точной, чем факторные модели.
Nagayasu (2014)	Инфляция	47 префектур Японии, 1976-2010 гг.	Между ценами регионов существует пространственная связь
Elhorst, Yesilyurt (2013)	Региональная инфляция	67 провинций Турции за период 1987–2001 гг.	Ожидаемая инфляция не влияет на фактическую инфляцию. Модель с пространственной корреляцией в ошибках не учитывает влияние побочных эффектов на инфляцию

3. Прогнозирование инфляции для европейских регионов России

При прогнозировании инфляции использовались данные по ИПЦ 49 регионов России, расположенных в Европейской части страны за 1999–2014 гг.

В качестве базовой модели (наивного прогноза), с которой будут сравниваться остальные прогнозы, использовались модели AR (1), оцененные для каждого из субъектов РФ отдельно с помощью метода наименьших квадратов. Для каждого региона были оценены модели вида

$$\pi_{it} = \alpha_i + \beta_i \pi_{it-1} + \varepsilon_{it}, \varepsilon_{it} \sim ND(0, \sigma_i^2), \quad (1)$$

Так как каждый ряд состоит из 15 наблюдений, МНК оценки оказываются смещенными, в связи с чем исследуются модели пула. При этом константа и коэффициент наклона предполагаются одинаковыми для всех временных и пространственных наблюдений.

$$\pi_{it} = \alpha + \beta \pi_{it-1} + \varepsilon_{it}, \varepsilon_{it} \sim NID(0, \sigma_i^2). \quad (2)$$

Данная модель оценивается двухшаговым обобщенным методом моментов на основании методики Ареллано-Бонда¹. Во всех оцененных регрессиях (здесь и далее) значимым оказался первый лаг инфляции плюс дополнительные лаги, разные для различных спецификаций.

Следующая используемая для прогнозирования модель – модель панельной регрессии с фиксированными эффектами, в которой константа предполагается регионально специфической переменной:

$$\pi_{it} = \alpha_i + \beta \pi_{it-i} + \varepsilon_{it}, \varepsilon_{it} \sim NID(0, \sigma_i^2) \quad (3)$$

Далее были построены регрессии, учитывающие пространственную корреляцию. Для оценки этих моделей также использовались данные, собранные в пул, и модели с фиксированными

¹ При использовании данного метода временные лаги зависимой переменной могут использоваться одновременно как в качестве объясняющих переменных, так и в качестве инструментов. При анализе инфляции использовались различные комбинации лаговых значений. Критерием качества служил тест Саргана, проверяющий сверхидентифицируемость инструментов (для оцененной регрессии нулевая гипотеза не была отвергнута; это верно и далее для всех моделей, оцененных данным методом).

эффектами. Все модели пространственной автокорреляции оценивались методом максимального правдоподобия. Для расчетов использовалась матрица весов на основе общей границы (w_{ij})¹: если регион i имеет общую морскую или сухопутную границу с регионом j , то $w_{ij} = 1$, в противном случае, $w_{ij} = 0$. Значения статистики I Морана оказались значимыми на всем временном промежутке.

Модель объединенной регрессии (пула) для всего периода имеет вид

$$\pi_{it} = \alpha + \beta \pi_{it-1} + \delta \sum_{j=1}^{49} w_{ij} \pi_{jt} + \varepsilon_{it}. \tag{4}$$

Далее была оценена модель с фиксированными эффектами с учетом пространственной корреляции

$$\pi_{it} = \alpha_i + \beta \pi_{it-1} + \delta \sum_{j=1}^{49} w_{ij} \pi_{jt} + \varepsilon_{it}. \tag{5}$$

Данные модели, учитывающие пространственную зависимость между регионами в лагах и в ошибках, оценивались методом максимального правдоподобия².

Функция правдоподобия для учета корреляции в лагах имеет вид

$$-\frac{NT}{2} \ln(2\pi\sigma^2) + T \sum_{i=1}^N \ln(1 - \delta w_i) - \frac{1}{2\sigma^2} \sum_{i=1}^T e_i' e_i, \quad e_i = (I - \delta W)(Y_i - \bar{Y}) - (X_i - \bar{X})\beta. \tag{6}$$

где w_i – характеристические корни весовой матрицы W ; σ^2 – дисперсия ошибок оцениваемого уравнения.

Функция правдоподобия для модели пространственной корреляции в ошибках принимает следующий вид:

$$-\frac{NT}{2} \ln(2\pi\sigma^2) + T \sum_{i=1}^N \ln(1 - \delta w_i) - \frac{1}{2\sigma^2} \sum_{i=1}^T e_i' e_i, \quad e_i = (I - \delta W)[Y_i - \bar{Y} - (X_i - \bar{X})\beta]. \tag{7}$$

Оценки коэффициентов, полученные по всем моделям, приведены в *табл. 2*.

Таблица 2

РЕЗУЛЬТАТЫ ОЦЕНКИ МОДЕЛЕЙ

	Без пространственных эффектов			С пространственными эффектами			
	I_{OLS}		P_{OLS}	F_{OLS}	P_{SAR}	F_{SAR}	F_{SEM}
	Модель (1)		Модель (2)	Модель (3)	Модель (4)	Модель (5)	Модель (6)
	Максимум	Минимум					
α	42,47	0,76	67,2		7,1		
β	0,61	-0,61	0,27	0,46	-0,01	0,04	-0,23
δ					0,97	0,3	0,57

На основе оцененных параметров каждой модели были построены внутривыборочные прогнозы с целью выявления наиболее точного метода прогнозирования. Значения средних абсолютных процентных ошибок прогнозов (МАРЕ), усредненные по всем регионам, приведены в *табл. 3*.

Полученные результаты не позволяют сделать однозначные выводы относительно того, какая модель наилучшим образом предсказывает инфляцию. В целом ошибки по всем моделям, кроме модели с пространственной корреляцией в ошибках, дают практически идентичные результаты. Для европейских регионов «наивная» AR(1) модель дает точные результаты, не уступающие по точности остальным моделям.

¹ Все диагональные элементы матрицы весов равны нулю.

² Для оценки таких моделей методом ММП был написан специальный программный код

Таблица 3

МАРЕ-МОДЕЛИ

Без пространственных эффектов			С пространственными эффектами		
AR(1)	P_{OLS}	F_{OLS}	P_{SAR}	F_{SAR}^*	F_{SEM}
Модель (1)	Модель (2)	Модель (3)	Модель (4)	Модель (5)	Модель (6)
2%	3%	2%	5%	5%	42%

* Коэффициент незначим.

Для пула с учетом пространственной корреляции и модели с фиксированными эффектами с учетом пространственной корреляции в ошибках коэффициент при первом лаге инфляции оказывается отрицательным. Наименьшее значение положительный коэффициент β принимает в модели с фиксированными эффектами с пространственными эффектами в лагах, в то время как коэффициент при взвешенной инфляции соседей для данной модели оказывается максимальным.

Таким образом, можно говорить, что введение пространственных эффектов в модель для регионов европейской части России довольно сильно ухудшает ее прогнозные свойства по сравнению с более простыми моделями. ●