

**КОИНТЕГРАЦИОННЫЙ АНАЛИЗ ПРЕДЛОЖЕНИЯ
НА РЫНКЕ ЖИЛЬЯ ДАЛЬНЕВОСТОЧНОГО
ФЕДЕРАЛЬНОГО ОКРУГА НА ОСНОВЕ
ПАНЕЛЬНОЙ МОДЕЛИ КОРРЕКЦИИ ОШИБОК****Шерышев Е. И. ORCID ID 0009-0001-0758-9503,
Михалева М. Ю. ORCID ID 0009-0006-1574-2631***Федеральное государственное образовательное бюджетное учреждение высшего образования
«Финансовый университет при Правительстве Российской Федерации»,
Москва, Российская Федерация, e-mail: mmikhaleva@fa.ru*

Применение стандартных методов анализа панельных данных – объединенной регрессии, моделей с фиксированными и случайными эффектами – в условиях нестационарности временных рядов и неоднородности регионов сопряжено с риском ложной регрессии, когда статистически значимые связи выявляются при их фактическом отсутствии. Указанная проблема приобретает особую актуальность для рынка жилья Дальневосточного федерального округа, где системные эконометрические исследования, учитывающие нестационарность и гетерогенность панельных данных, до настоящего времени не проводились. Цель работы состоит в количественной оценке краткосрочных и долгосрочных взаимосвязей между предложением на рынке жилья, инвестициями в основной капитал и оборотом розничной торговли на основе панельной модели коррекции ошибок. Эмпирическая база включает сбалансированную панель по 11 субъектам Дальневосточного федерального округа за 2010–2023 гг.; спецификация модели является логарифмически линейной. На основе спецификационных тестов (на выбор между фиксированными и случайными эффектами, на наличие общих ненаблюдаемых факторов и на гомогенность коэффициентов наклона) обосновано применение метода оценки средних по группам с коррекцией на общие факторы, позволяющего учитывать как гетерогенность параметров, так и воздействие общих ненаблюдаемых эффектов. На основе методов обучения без учителя выявлен кластер регионов, наиболее похожих друг на друга. В результате получены оценки краткосрочных эластичностей предложения по инвестициям и обороту розничной торговли, а также оценена скорость возвращения рынка к долгосрочному равновесию после краткосрочных отклонений. Разработанная модель может служить инструментом анализа и прогнозирования предложения на рынке жилья Дальневосточного федерального округа.

Ключевые слова: Дальневосточный федеральный округ, рынок жилья, коинтегрированные временные ряды, модель коррекции ошибок***Благодарность:** Авторы выражают благодарность федеральному автономному научному учреждению «Восточный центр государственного планирования» за поддержку при проведении исследования.***COINTEGRATION ANALYSIS OF HOUSING SUPPLY
IN THE FAR EASTERN FEDERAL DISTRICT BASED
ON A PANEL ERROR CORRECTION MODEL****Sheryshev E. I. ORCID ID 0009-0001-0758-9503,
Mikhaleva M. Yu. ORCID ID 0009-0006-1574-2631***Federal State Educational Budgetary Institution of Higher Education
“Financial University under the Government of the Russian Federation”,
Moscow, Russian Federation, e-mail: mmikhaleva@fa.ru*

The application of standard panel data methods – pooled regression, fixed effects and random effects models – in the presence of non-stationary time series and regional heterogeneity entails the risk of spurious regression, where statistically significant relationships are detected despite their actual absence. This issue is particularly relevant for the housing market of the Far Eastern Federal District, where systematic econometric studies accounting for non-stationarity and heterogeneity of panel data have not been conducted to date. The aim of the study is to quantitatively assess the short-term and long-term relationships between the housing market supply, fixed capital investment, and retail trade turnover using a panel error correction model. The empirical basis comprises a balanced panel of eleven subjects of the Far Eastern Federal District for the period 2010–2023; the model specification is log-linear. Based on specification tests – for the choice between fixed and random effects, for the presence of common unobserved factors, and for slope homogeneity – the use of the Common Correlated Effects Mean Group estimator is justified, as it accounts for both parameter heterogeneity and the impact of common unobserved effects. Unsupervised learning algorithms have been used to find a cluster of regions that share similar properties. Estimates of short-run elasticities of housing supply with respect to investment and retail trade turnover are obtained, along with an estimate of the speed of adjustment of the market to long-run equilibrium following short-run deviations. The developed model can serve as a tool for analysis and forecasting of housing supply in the Far Eastern Federal District.

Keywords: Far Eastern Federal District, housing market, cointegrated time series, error correction model***Acknowledgments:** The authors express their gratitude to the Federal Autonomous Scientific Institution “Eastern Center for State Planning” for its support in conducting the research.*

Введение

Сегодня рынок жилья является объектом повышенного интереса со стороны всех агентов экономики в силу своего стремительного развития. С развитием рынка жилья, соответственно, возрастает интерес к исследованиям, посвященным прогнозированию его состояния, выявлению закономерностей и взаимосвязей на нем [1–3]. В то время как продвинутые модели машинного обучения часто работают как «черный ящик» и ориентированы исключительно на прогноз, эконометрические методы позволяют не только прогнозировать, но и содержательно интерпретировать количественные взаимосвязи между факторами, а также проверять статистические гипотезы об их значимости [4–6].

В доступной литературе по эконометрическому моделированию региональных рынков жилья преобладают работы, использующие стандартные подходы к панельным данным [7–9]. Вместе с тем, согласно [10, 11], учет нестационарности временных рядов в подобных исследованиях не проводится. В доступных работах, непосредственно посвященных рынку жилья Дальневосточного федерального округа, подобные методологические подходы не обнаружены.

Стандартные подходы к работе с панельными данными, объединенная модель, модели с фиксированными и случайными эффектами, опираются на гомогенность наклонных коэффициентов и стационарность индивидуальных временных рядов. При нарушении этих условий стандартные методы могут не только не выявлять существующие долгосрочные связи, но и давать статистически значимые оценки при их фактическом отсутствии (феномен ложной регрессии).

В силу подобных ограничений исследования нестационарных панельных данных с коинтеграционными связями могут проводиться на основе модели коррекции ошибок (Error Correction Model, ECM). Однако следует отметить, что состоятельность оценок ECM зависит от корректного учета структуры панели. В условиях гетерогенности коэффициентов и наличия общих ненаблюдаемых факторов предпочтительными являются подходы, основанные на оценивании средних по группе (Mean Group, MG) или оценивании средних, скорректированных на общие факторы по группе (Common Correlated Effects Mean Group, CCEMG) [12]. В настоящей работе используется подход CCEMG, позволяющий получить состоятельные оценки краткосрочных эластичностей и скорости корректировки.

Цель исследования – количественная оценка краткосрочных и долгосрочных взаимосвязей между предложением на рынке жилья, инвестициями в основной капитал и оборотом розничной торговли в 11 регионах Дальневосточного федерального округа на основе панельной модели коррекции ошибок.

Материалы и методы исследования

Для исследования использованы данные по 11 регионам Дальневосточного федерального округа за период с 2010 по 2023 г. (включительно). Таким образом, структура панели сбалансирована и составляет $N = 11$, $T = 14$, где N – число регионов, T – число временных периодов.

Исходные статистические данные предоставлены федеральным автономным научным учреждением «Восточный центр государственного планирования» в рамках научно-исследовательской работы по теме «Эконометрический анализ рынка жилья (на примере Дальневосточного федерального округа)». Первичными источниками данных являются официальные публикации Федеральной службы государственной статистики (Росстат), включая показатели ввода жилья, инвестиций в основной капитал и оборота розничной торговли по всем субъектам ДФО за 2010–2023 гг. В связи с изменением административного состава ДФО (включение Республики Бурятия и Забайкальского края с 2018 г.) в работе используется текущая структура округа из 11 регионов; данные за период до 2018 г. пересчитаны в соответствии с современными границами. Упомянутый пересчет был проведен организацией, предоставившей данные. Единицы измерения: ввод жилья – тыс. м², инвестиции и оборот розницы – млрд руб.

В качестве эндогенной переменной в работе рассматривается показатель ввода общей площади жилых помещений (с учетом жилых домов на участках для ведения садоводства), тыс. м². В настоящей работе под предложением на региональном рынке жилья понимается именно этот показатель. Переход от физического ввода к рыночному предложению обоснован тем, что ввод является количественным измерителем объема новой жилой площади, поступающей на рынок. Вместе с тем признаются ограничения: показатель не охватывает вторичный рынок и не дифференцирует жилье по целевому назначению, что следует учитывать при интерпретации результатов.

Экзогенные переменные – объем инвестиций в основной капитал, млрд руб., и оборот розничной торговли, млрд руб.

Для устранения влияния инфляции все стоимостные показатели приведены к сопоставимым ценам с использованием соответствующих дефляторов, публикуемых Росстатом.

В работе используется логарифмически линейная спецификация модели: все переменные преобразованы с помощью натурального логарифма. Данный выбор обусловлен, с одной стороны, возможностью интерпретировать коэффициенты модели как эластичности, а с другой – необходимостью стабилизации дисперсии и приведения переменных к единому масштабу.

Спецификация модели коррекции ошибок первого порядка имеет вид

$$\Delta y_t = \beta_0 + \beta_1 \cdot ECT + \beta_2 \cdot \Delta x_{1t} + \dots + \beta_{n+1} \cdot \Delta x_{nt} + \varepsilon_t,$$

где

$$ECT = y_{t-1} - \alpha_0 - \alpha_1 \cdot x_{1(t-1)} - \dots - \alpha_n \cdot x_{n(t-1)}$$

$$-1 < \beta_1 < 0.$$

Для обоснования применения модели коррекции ошибок и выбора спецификации долгосрочного уравнения был реализован ряд статистических процедур. На первом этапе с помощью теста Хаусмана определен вид эффектов в долгосрочном уравнении, а F -тест использован для проверки гомогенности наклонных коэффициентов. Стационарность индивидуальных рядов проверялась тестом Зивота – Эндрюса, допускающим наличие структурных сдвигов, а также тестом CIPS (Cross-Sectionally Augmented Im-Pesaran-Shin), учитывающим возможную кросс-секционную зависимость в панели [13] в качестве дополнительного инструмента (с учетом его ограниченной мощности при $T = 14$). Наличие коинтеграции между переменными тестировалось с помощью теста Вестерлунда [14], кросс-секционная зависимость – с помощью CD-теста (Cross-Sectional Dependence test) [15]. Оценка краткосрочных параметров выполнена с использованием трех подходов: MG (Mean Group) – коэффициенты рассчитываются как средние по всем панелям; CСЕР (Common Correlated Effects Pooled) – предполагает гомогенность коэффициентов и дополняет уравнение панельными средними для фильтрации общих ненаблюдаемых факторов; CСЕМГ (Common Correlated Effects Mean Group) – допускает гетерогенность коэффициентов и также включает панельные средние.

Для кластеризации регионов и визуализации кластеров использовались методы

обучения без учителя HDBSCAN, позволяющий разделить данные на кластеры и аномалии, и t-SNE, позволяющий преобразовать многомерное пространство признаков кластеризации в двухмерное пространство, в котором получившиеся кластеры можно визуализировать.

В работе принят 5% уровень значимости для отвержения нулевой гипотезы. При этом коэффициенты, статистическая значимость которых незначительно превышает 10% порог, не исключаются из моделей ввиду ограниченного объема выборки – их удаление сопряжено с риском потери важного фактора. Панельные средние сохраняются в уравнениях даже при отсутствии статистической значимости.

Для проверки адекватности моделей выполнены тесты на наличие единичного корня в остатках. Нестационарность остатков свидетельствует о непригодности модели. Расчеты выполнены на языке программирования Python с использованием библиотек statsmodels, sklearn, seaborn и Westerlund. Для методов, не реализованных в существующих библиотеках, авторами разработана собственная библиотека econmethods [репозиторий]. В качестве основной метрики оценки предиктивной способности модели использована метрика MASE (mean absolute scaled error).

Результаты исследования и их обсуждение

Судя по всему спектру описательных статистик (табл. 2–4) и графикам исходных временных рядов (рис. 1–3), регионы немало отличаются друг от друга. Именно поэтому все переменные логарифмированы, что приводит все числовые статистики к более благоприятной для оценки модели шкале. Во временных рядах розничной торговли в 2018 г. замечен, с одной стороны, выброс, с другой стороны, вероятно, объяснимое явление, ведь этот «выброс» присутствует у всех рядов, поэтому в настоящей работе экзогенная переменная x_2 не подвергалась никаким преобразованиям, кроме логарифмирования. В табл. 1 показан фрагмент датасета, используемого для оценки моделей.

Для проверки стационарности применены тесты Зивота – Эндрюса и CIPS (последний – с учетом ограниченной мощности при $T = 14$). Результаты представлены в табл. 5, 6. Тест Зивота – Эндрюса указывает на нестационарность всех рядов; тест CIPS дает аналогичный вывод для предложения и торговли, однако отвергает нестационарность для инвестиций (табл. 6).

Таблица 1

Фрагмент датасета с логарифмированными переменными

region	year	target	x1	x2
Амурская область	2010	5,113793386	4,429531819	4,203716023
Амурская область	2011	5,48438191	4,713982731	4,395185435
Амурская область	2012	5,760225818	4,605027506	4,523278068
'''	'''	'''	'''	'''
Забайкальский край	2015	5,459585514	3,919018781	4,574613665
Забайкальский край	2016	5,675725877	3,982879203	4,567305467
Забайкальский край	2017	5,600272098	4,075040161	4,571007983
Забайкальский край	2018	5,228431239	3,965141627	1,956589591
'''	'''	'''	'''	'''
Чукотский автономный округ	2021	1,106249716	3,234120599	1,726168968
Чукотский автономный округ	2022	1,25874504	3,593852806	1,799403121
Чукотский автономный округ	2023	1,826321914	3,588958403	1,794668502

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 2

Описательная статистика ввода жилья

region	target				
	mean	median	std	max	min
Амурская область	265,4684286	233,601	100,9398928	496,203	156,1
Еврейская автономная область	47,07471429	51,5	24,58608719	104	14,1
Забайкальский край	267,9309286	275,85	58,6241255	380,245	180,4
Камчатский край	67,56435714	68,45	14,70902667	87,7	35,5
Магаданская область	12,12064286	11,7425	5,356580017	20,7	4,3
Приморский край	665,1752143	601,315	226,1894196	1209,115	412,1
Республика Бурятия	325,5036429	298,085	68,26027984	469,478	247,8
Республика Саха (Якутия)	511,353	538,15	121,9605596	705,128	303
Сахалинская область	342,9942857	321,8	106,8442267	512,88	201,8
Хабаровский край	357,1171429	339,202	90,09166333	577,797	229,3
Чукотский автономный округ	2,295357143	1,95	1,553249428	6,211	0,3
Все регионы	260,417974	268,7	225,454739	1209,115	0,3

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 3

Описательная статистика инвестиций в основной капитал

region	x1				
	mean	median	std	max	min
Амурская область	141,0204709	114,5456681	73,50826958	310,2811303	56,8060238
Еврейская автономная область	11,84667594	9,18833898	5,625717611	23,9676164	6,746470337
Забайкальский край	63,16143425	54,18664785	21,94797959	127,0638838	44,824818
Камчатский край	30,32775137	29,4078803	10,56918129	56,57562011	15,0823868
Магаданская область	29,1615872	28,55238419	7,3398135	40,04378482	16,808723
Приморский край	134,5985028	113,9930887	59,65034524	289,9318106	79,99911346
Республика Бурятия	35,83657193	35,0924258	12,28528096	69,69664558	20,94514028
Республика Саха (Якутия)	196,3860203	180,4496076	58,54190449	310,2856667	130,492937
Сахалинская область	139,0817316	135,2949271	14,39828849	165,9006946	120,4067626
Хабаровский край	120,0553329	123,9085918	37,20128716	183,837398	74,04228203
Чукотский автономный округ	15,16412457	10,40600469	10,37672583	36,37394806	5,418724
Все регионы	83,330928	56,690822	71,076468	310,285667	5,418724

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 4

Описательная статистика оборота розничной торговли

region	x2				
	mean	median	std	max	min
Амурская область	91,01562388	98,47283149	26,73929091	111,3589146	7,445902408
Еврейская автономная область	13,74817699	14,69638745	3,683124513	15,76344265	1,083064847
Забайкальский край	92,58421332	98,79599101	24,82850831	105,3652968	7,075156698
Камчатский край	33,09269903	34,72106827	9,136930791	42,11478275	2,618065725
Магаданская область	17,30655208	18,75106917	4,83852433	20,49372011	1,453693398
Приморский край	209,9350649	223,6284634	63,04914661	269,0633844	15,28317029
Республика Бурятия	97,21271379	105,9483243	26,61765114	112,193441	7,822037373
Республика Саха (Якутия)	119,2729166	125,9445529	34,09368661	147,2199517	8,031105383
Сахалинская область	83,2299248	88,90379014	22,13244242	93,53798257	6,840665456
Хабаровский край	168,4029776	180,7749735	47,26189999	197,1674257	13,75578466
Чукотский автономный округ	4,943718371	5,593437285	1,491892942	6,046037635	0,431023373
Все регионы	84,613144	91,603705	69,015403	269,063384	0,431023

Примечание: составлена авторами на основе полученных данных в ходе исследования.

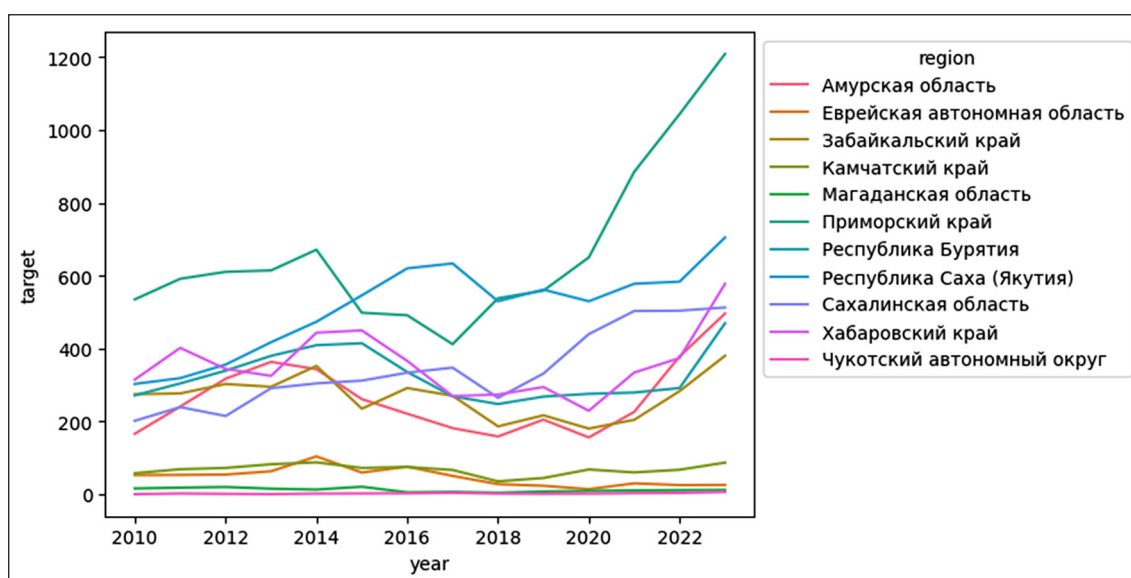


Рис. 1. Временные ряды эндогенной переменной «Ввод жилья» в региональном разрезе

Примечание: составлен авторами по результатам данного исследования

Таблица 5

Тест Зивота – Эндрюса

	Предложение	Инвестиции	Торговля
Статистика теста	-4,2483	-3,797	-3,88
Критическое значение	-4,81067	-4,81067	-4,81067
P-Value	0,215	0,48	0,426
Параметры	Const	Const	Const
Вывод	Выявлена нестационарность	Выявлена нестационарность	Выявлена нестационарность

Примечание: составлена авторами на основе полученных данных в ходе исследования.

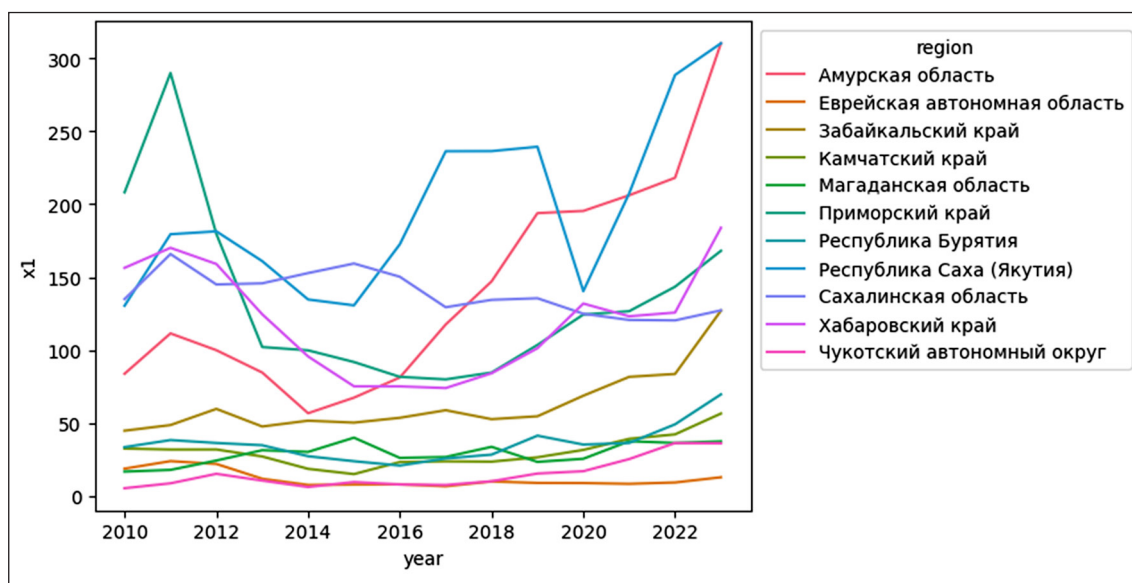


Рис. 2. Временные ряды экзогенной переменной
«Инвестиции в основной капитал» в региональном разрезе
Примечание: составлен авторами по результатам данного исследования

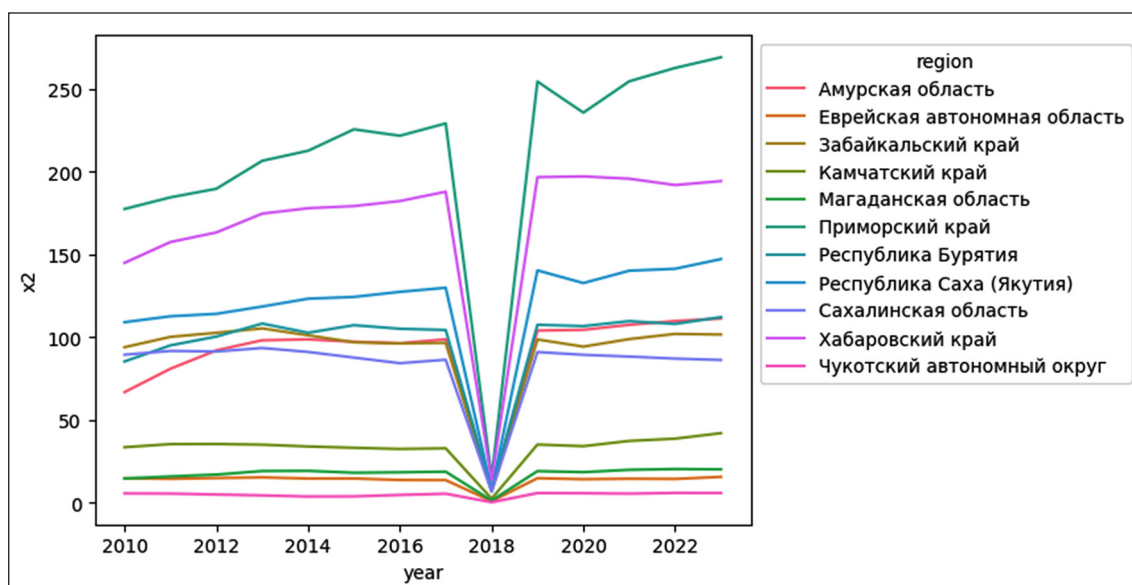


Рис. 3. Временные ряды экзогенной переменной
«Оборот розничной торговли» в региональном разрезе
Примечание: составлен авторами по результатам данного исследования

Таблица 6

Тесты CIPS

	Предложение	Инвестиции	Торговля
Статистика теста	-0,779	-3,031	-1,281
Критическое Значение	-2,37	-2,37	-2,37
Параметры	Const	Const	Const
Вывод	Выявлена нестационарность	Выявлена стационарность	Выявлена нестационарность

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 7

Тест Вестерлунда

Статистика	Gt	Ga	Pt	Pa
P-Value	0,0086	0,176	0,0034	0,154

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 8

Тесты на гомогенность наклонных коэффициентов,
панельную созависимость и тест Хаусмана

	P-Value	Вывод
F-тест на гомогенность	1,88e-41	H1 (коэффициенты гетерогенны)
CD-тест	5,733e-16	H1 (Найдены общие ненаблюдаемые факторы)
Тест Хаусмана	1,0	H0 (Случайные эффекты)

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 9

Долгосрочная модель

Регрессор/метрика	Коэффициент	P-Value / значение
x1	0,7914	0,004
x2	0,2919	0,007
R ²	–	0,776
F-test	–	3,88e-50
R ² adj	–	0,773

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 10

Оценки ECM

Регрессор/Метод	MG	ССЕР	ССЕМГ
Δx_1	0,2	0,0647	0,428
Δx_2	0,08	0,121	0,196
$\bar{y}_t$	–	-0,309	0,537
$\bar{x}_{1(t)}$	–	0,246	-0,433
$\bar{x}_{2(t)}$	–	0,13	-0,168
ECT	-0,03	-0,08	-0,474

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Наличие коинтеграции между переменными проверялось с помощью теста Вестерлунда (табл. 7). Вывод основан на статистиках Pt и Gt , рассчитываемых на основе t -статистик коэффициентов при коэффициенте коррекции ошибок (Error Correction Term, ECT). Поскольку отсутствие коинтеграции исключает корректное применение ЕСМ, подтверждение коинтеграции является необходимым условием для построения модели. Тест Вестерлунда подтверждает наличие коинтеграции.

Для определения спецификации модели коррекции ошибок проведены тест Хаусмана, CD-тест и F-тест на гомогенность наклонных коэффициентов. Результаты представлены в табл. 8. Тест Хаусмана указывает на предпочтительность модели со случайными эффектами; CD-тест свидетельствует о наличии общих ненаблюдаемых факторов; F-тест подтверждает гетерогенность наклонных коэффициентов.

По результатам проведенного тестирования оценены: долгосрочная модель

равновесия со случайными эффектами (табл. 9), ЕСМ с парадигмами оценивания MG, ССЕР, ССЕМГ (табл. 10, 11). Для ССЕР панельные средние лагированы на год, в отличие от ССЕМГ.

В табл. 9–11 x_1 – логарифм инвестиций в основной капитал, x_2 – логарифм оборота розничной торговли, ECT – член коррекции ошибок.

На основе проведенных расчетов оценены три модели коррекции ошибок:

Mean Group:

$$\Delta y_t = -0.03 \cdot ECT + 0.2 \cdot \Delta x_{1t} + 0.08 \cdot \Delta x_{2t} + \varepsilon_t. \quad (1)$$

Common Correlated Effects Pooled:

$$\Delta y_t = -0.08 \cdot ECT + 0.121 \cdot \Delta x_{2t} - 0.309 \cdot \bar{y}_{t-1} + 0.246 \cdot \bar{x}_{1(t-1)} + 0.13 \cdot \bar{x}_{2(t-1)} + \varepsilon_t. \quad (2)$$

Common Correlated Effects Mean Group:

$$\Delta y_t = -0.474 \cdot ECT + 0.428 \cdot \Delta x_{1t} + 0.196 \cdot \Delta x_{2t} + 0.537 \cdot \bar{y}_t - 0.433 \cdot \bar{x}_{1(t)} - 0.168 \cdot \bar{x}_{2(t)} + \varepsilon_t. \quad (3)$$

Таблица 11

Статистики ЕСМ

Статистика/Метод	MG	ССЕР	ССЕМГ
Δx_1 (p-value)	0,168	0,592	0,03
Δx_2 (p-value)	0,002	0,004	0,002
$\bar{y}_t$ (p-value)	–	0,006	0,06
$\bar{x}_{1(t)}$ (p-value)	–	0,04	0,1
$\bar{x}_{2(t)}$ (p-value)	–	0,049	0,04
ECT (p-value)	0,25	0,003	0,001
F-тест (p-value)	–	5,5e-06	–
W-тест (p-value)	0,004	–	5,325e-13
R ²	0,295	0,206	0,72

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 12

Тест Зивота – Эндрюса на остатках

Статистика / Метод	MG	ССЕР	ССЕМГ
Критическое значение (5 %)	-4,81067	-4,81067	-4,81067
Статистика	-7,28	-12,68	-4,47
P-Value	1e-05	1e-05	0,13
Параметры	Const	Const	Const
Вывод	Выявлена стационарность	Выявлена стационарность	Выявлена нестационарность

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 13

Тест CIPS на остатках

Статистика / Метод	MG	ССЕР	ССЕМГ
Критическое значение (5 %)	-1,74	-1,74	-2,37
Статистика	-2,36	-2,39	-1,76
Параметры	–	–	Const
Вывод	Выявлена стационарность	Выявлена стационарность	Выявлена нестационарность

Примечание: составлена авторами на основе полученных данных в ходе исследования.

В спецификациях ССЕР используются также лагированные панельные средние. Как показано в [10], на малых выборках включение средних в текущий период может приводить к поглощению эффектов индивидуальных регрессоров, тогда как добавление лаговых значений позволяет более полно фильтровать влияние общих ненаблюдаемых факторов. Однако спецификация CCEMG, в отличие от ССЕР, показывает лучший результат при нелагированных на один период средних. Среди трех оцененных спецификаций наивысший коэффициент детерминации получен для модели (3) ($R^2 = 0,72$). Модель (1) характеризуется $R^2 = 0,295$, модель (2) – $R^2 = 0,206$.

На заключительном этапе остатки моделей (1)–(3) проверены на стационарность с использованием тестов Зивота – Эндрюса и CIPS (табл. 12, 13), проверены метрики качества MAE, MedAE, RMSE, MASE на выборках hold-three-out, hold-two-out. Нестационарность остатков служит основанием для признания модели неприменимой. Результаты тестирования свидетельствуют о стационарности остатков моделей (1) и (2); в то же время оба теста не позволяют отвергнуть гипотезу о наличии единичного корня в остатках модели (3).

Проведена оценка качества модели MG (1), ССЕР (2), CCEMG (3) на выборке hold-three-out: (2010–2021) в train / (2021–2023) в test, выборке hold-two-out: (2010–2022)

в train / (2022–2023) в test (табл. 14 и 15 соответственно).

На рис. 4–6 представлены истинные и предсказанные первые разности по всем моделям на выборке hold-three-out.

Таблица 14

Таблица метрик hold-three-out

Метрика / Модель	MG	ССЕР	CCEMG
MAE	0,194	0,164	0,427
RMSE	0,254	0,203	0,485
MedAE	0,163	0,151	0,452
MASE	0,79	0,67	1,74

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 15

Таблица метрик hold-two-out

Метрика/ Модель	MG	ССЕР	CCEMG
MAE	0,24	0,187	0,426
RMSE	0,293	0,221	0,477
MedAE	0,234	0,158	0,454
MASE	1	0,79	1,79

Примечание: составлена авторами на основе полученных данных в ходе исследования.

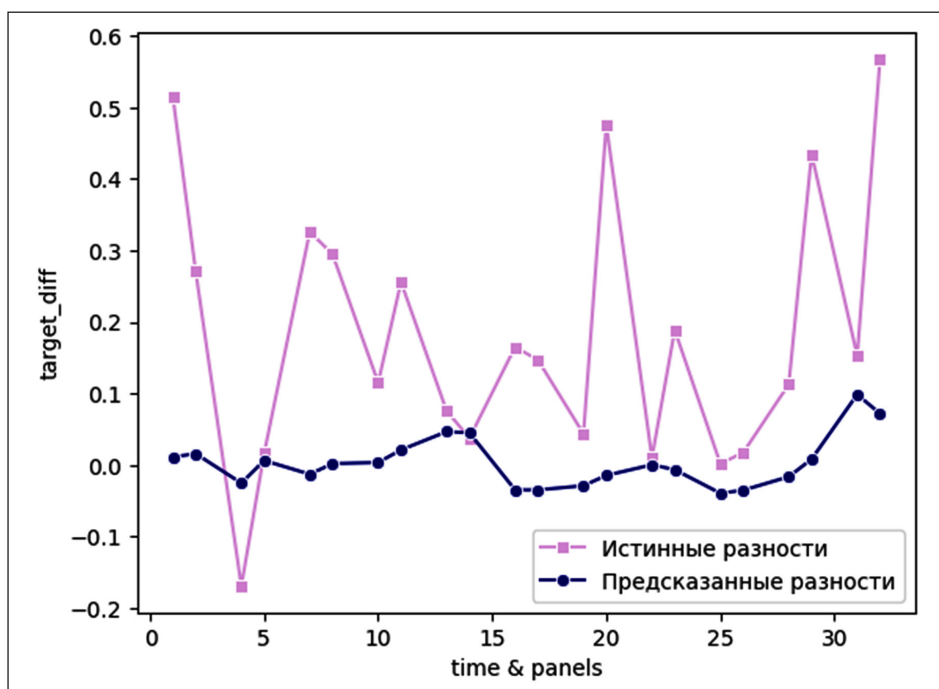


Рис. 4. Истинные и предсказанные разности модели MG (1) на разбиении hold-three-out
Примечание: составлен авторами по результатам данного исследования

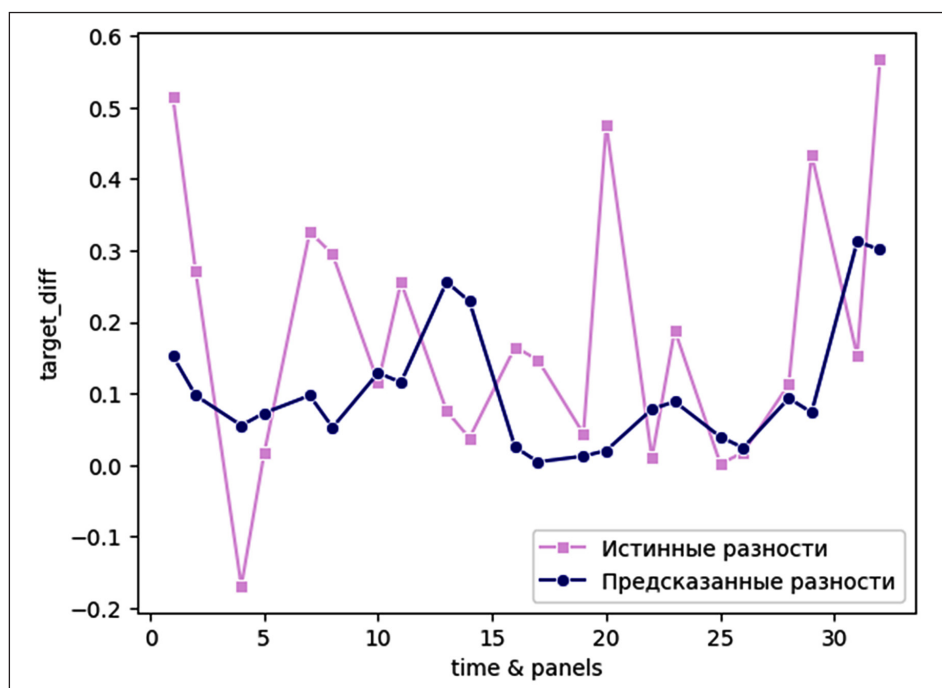


Рис. 5. Истинные и предсказанные различия модели CSEP (2) на разбиении hold-three-out
Примечание: составлен авторами по результатам данного исследования

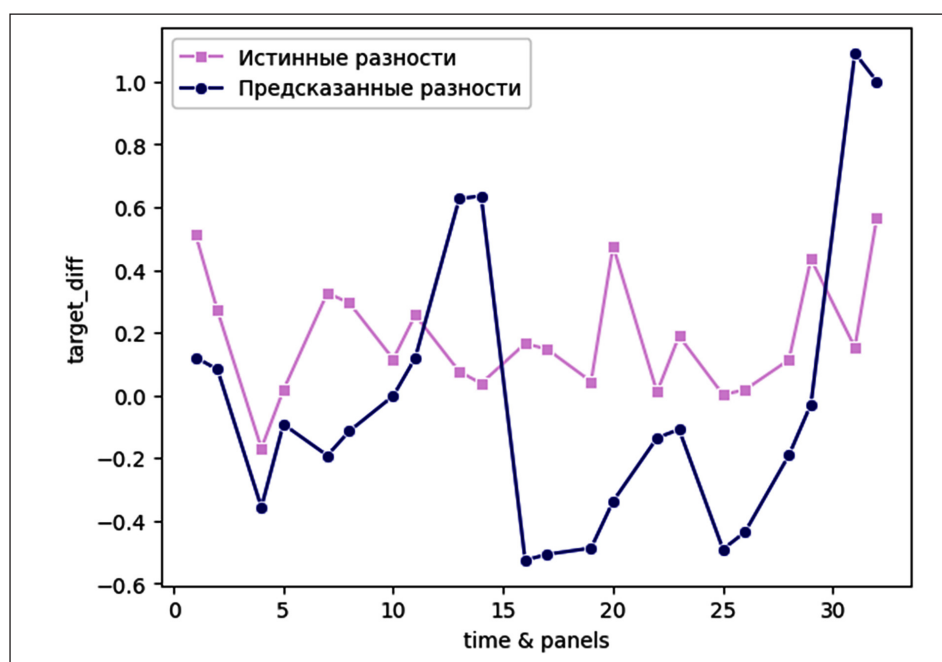


Рис. 6. Истинные и предсказанные различия модели CCEMG (3) на разбиении hold-three-out
Примечание: составлен авторами по результатам данного исследования

В ходе основного исследования было замечено, что некоторые регионы Дальневосточного федерального округа значительно отклоняются по поведению первых разностей зависимой переменной от остальных. Учитывая это наблюдение и тот

факт, что модель (3) приспосабливается к аномальным регионам, дополнительно была оценена модель CCEMG на подвыборке из восьми регионов (табл. 17–19), составляющих более гомогенный кластер, чем все регионы сразу, проведена оценка

ее эффективности (табл. 21) и стационарности остатков (табл. 20). Отбор регионов и визуализация кластеров проводились с помощью методов обучения без учителя HDBSCAN и t-SNE соответственно. Методом HDBSCAN обнаружено два мини-кластера и три аномальных региона (рис. 7). Разделение датасета на мини-кластеры в условиях кросс-секциональной зависимости может привести к зашумлению панельных

средних, так как оценка модели на подвыборке подразумевает полную изоляцию регионов подвыборки от регионов-девиантов, поэтому упомянутые мини-кластеры были объединены в один суперкластер из восьми регионов (рис. 8). В качестве признаков для кластеризации использовались: стандартное отклонение, коэффициент асимметрии, эксцесс и медиана первых разностей эндогенной переменной (табл. 16).

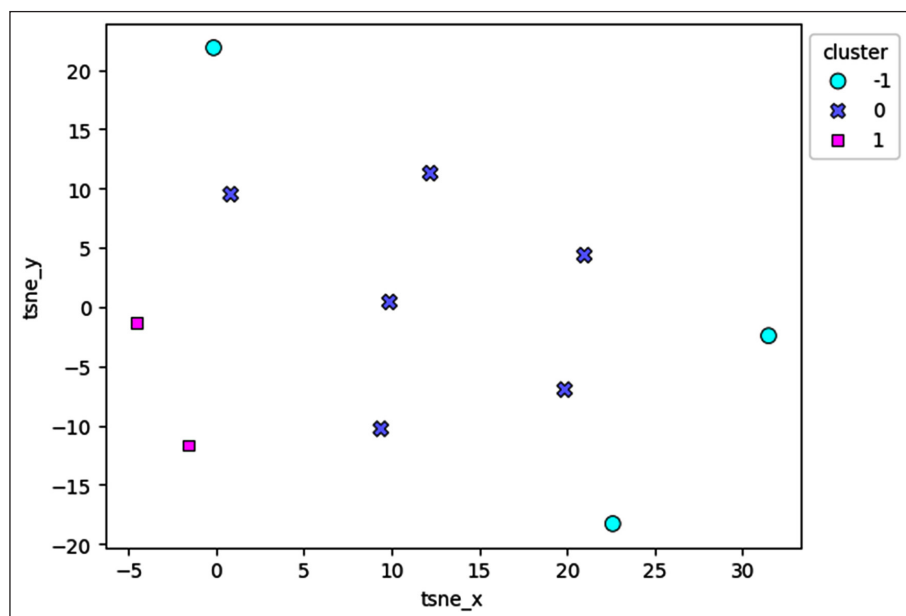


Рис. 7. Визуализация кластеризации HDBSCAN на мини-кластерах с помощью t-SNE.

Точки с индексом -1 определены как аномалии

Примечание: составлен авторами по результатам данного исследования

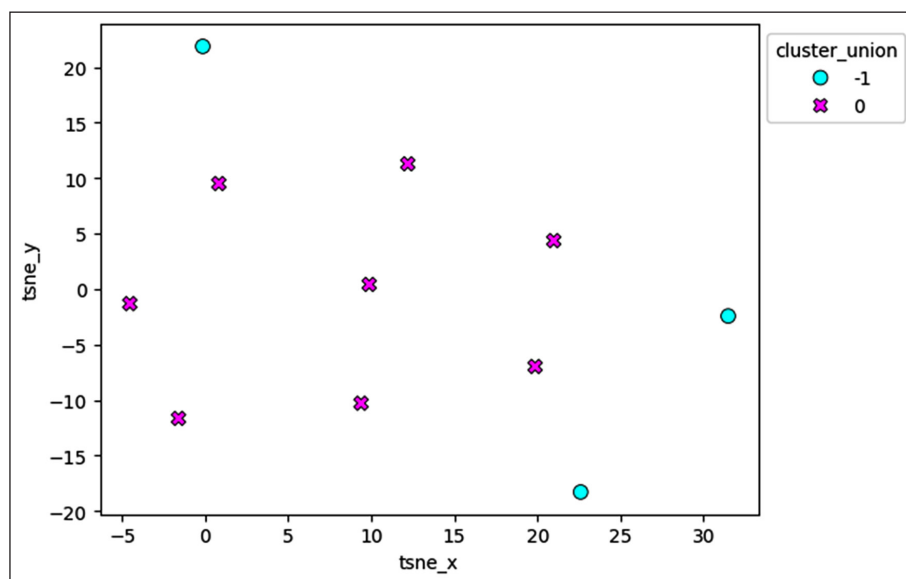


Рис. 8. Визуализация кластеризации HDBSCAN на суперкластере с помощью t-SNE.

Точки с индексом -1 определены как аномалии

Примечание: составлен авторами по результатам данного исследования

Таблица 16

Признаки кластеризации регионов и метки кластеров

Region	Медиана	Ст. отклонение	Асимметрия	Эксцесс	Кластер	Супер-кластер
Амурская область	0,14	0,28	0,02	-1,65	0	0
Еврейская автономная область	0,01	0,41	0,45	-0,29	1	0
Забайкальский край	0,09	0,23	-0,71	-0,34	0	0
Камчатский край	0,06	0,26	-1,26	2,74	0	0
Магаданская область	0,1	0,48	-1,92	5,01	-1	-1
Приморский край	0,09	0,17	-0,78	0,86	0	0
Республика Бурятия	0,04	0,17	0,89	3,06	-1	-1
Республика Саха (Якутия)	0,09	0,1	-1,26	1,79	0	0
Сахалинская область	0,04	0,16	-0,52	0,69	0	0
Хабаровский край	0,02	0,24	0,18	-1,12	1	0
Чукотский автономный округ	0,2	0,85	0,48	0,29	-1	-1

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 17

Долгосрочная модель на восьми регионах

Регрессор/метрика	Коэффициент	P-Value / значение
x1	0,9067	0
x2	0,3027	0
R ²	—	0,944
F-test	—	8,77e-70
R ² adj	—	0,943

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 19

Статистики ECM на восьми регионах

Статистика/Метод	CCEMG
Δx_1 (p-value)	0,0768
Δx_2 (p-value)	0,015
$\bar{y}_t$ (p-value)	0,23
$\bar{x}_{1(t)}$ (p-value)	0,27
$\bar{x}_{2(t)}$ (p-value)	0,28
ECT (p-value)	0,019
W-тест (p-value)	6,45e-07
R ²	0,672

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 18

Оценки ECM на восьми регионах

Регрессор/Метод	CCEMG
Δx_1	0,293
Δx_2	0,13
$\bar{y}_t$	0,357
$\bar{x}_{1(t)}$	-0,333
$\bar{x}_{2(t)}$	-0,079
ECT	-0,355

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Асимметрия, стандартное отклонение и эксцесс позволяют всецело описать неоднозначность распределений первых разностей зависимой переменной каждого региона. Медиана вносит информацию про типичное поведение первых разностей за счет устойчивости к экстремальным значениям.

Гиперпараметр, задающий минимальное количество элементов в кластере, для метода HDBSCAN равен 2 (`min_cluster_size = 2`) в силу ограниченности выборки в 11 элементов.

Гиперпараметр `perplexity` (эффективное количество соседей у каждой точки) для метода t-SNE выбран на основе числа неаномальных регионов + 1 (`perplexity = 9`), так как это позволяет визуализировать случай наибольшей близости кластеров и дает алгоритму шанс рассматривать один из аномальных регионов как близкого соседа для каждой точки. Также рассчитана метрика качества разделения кластеров – `silhouette score`. Для кластеров 1 и 0 она составила 0,39. Так как точки с меткой -1 не считаются за кластер, между ними и остальными кластерами `silhouette score` не рассчитывается.

Подвыборка содержит данные по восьми регионам Дальневосточного федерального округа за период с 2010 по 2023 г. (включительно). Панель сбалансирована и составляет $N = 8$, $T = 14$. В подвыборку не вошли: Чукотский автономный округ, Магаданская область, Республика Бурятия.

Полученная модель Common Correlated Effects Mean Group для восьми регионов:

$$\Delta y_t = -0.355 \cdot \text{ECT} + 0.293 \cdot \Delta x_{1t} + 0.13 \cdot \Delta x_{2t} + 0.357 \cdot \bar{y}_t - 0.333 \cdot \bar{x}_{1(t)} - 0.079 \cdot \bar{x}_{2(t)} + \varepsilon_t. \quad (4)$$

Таблица 20

Тесты на остатках модели (4)

Статистика/Метод	Зивота – Эндрюса	CIPS
Критическое значение (5 %)	-4,81067	-1,74
Статистика	-5,25	-1,88
P-value	0,012	–
Параметры	Const	–
Вывод	Выявлена стационарность	Выявлена стационарность

Примечание: составлена авторами на основе полученных данных в ходе исследования.

Таблица 21

Таблица метрик для модели (4)

Метрика/Выборка	Hold-Three-Out	Hold-Two-Out
MAE	0,12	0,11
RMSE	0,136	0,13
MedAE	0,11	0,058
MASE	0,62	0,55

Примечание: составлена авторами на основе полученных данных в ходе исследования.

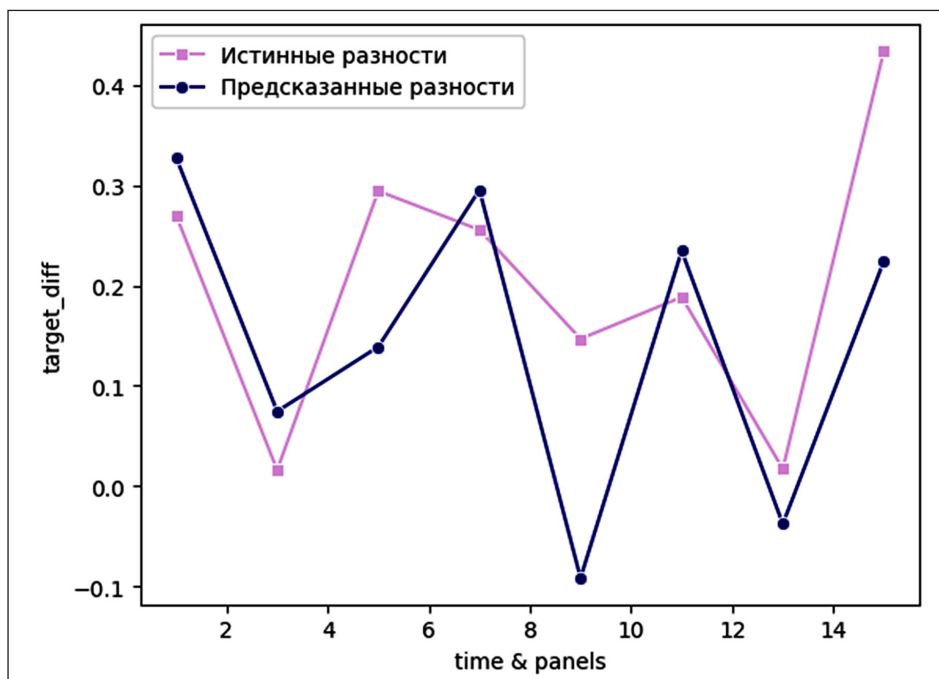


Рис. 9. Истинные и предсказанные различия модели CCEMG (4) на разбиении hold-three-out
Примечание: составлен авторами по результатам данного исследования

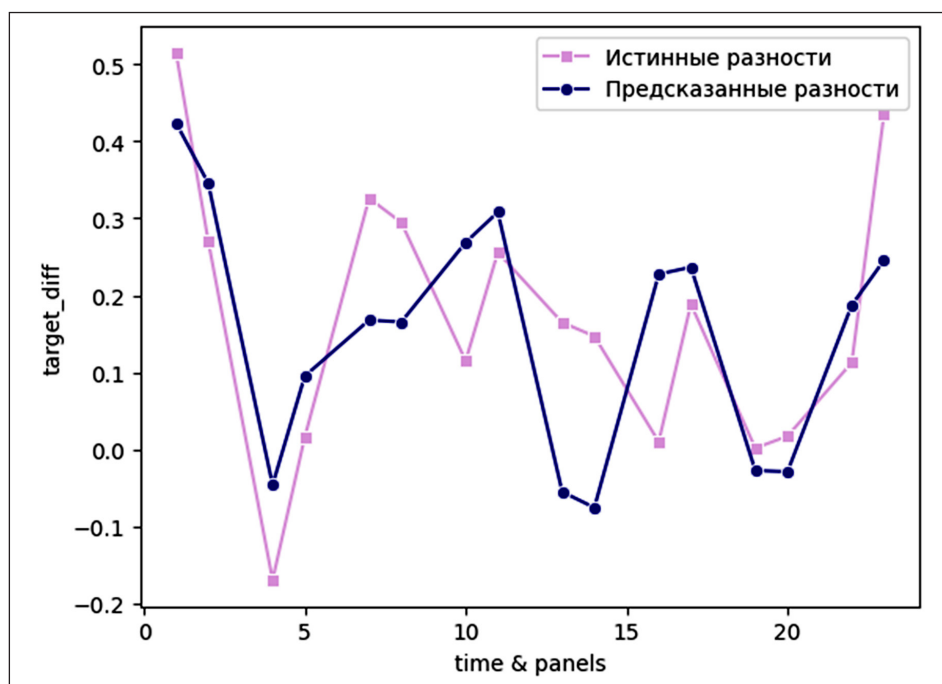


Рис. 10. Истинные и предсказанные разности модели CCEMG (4) на разбиении hold-two-out
Примечание: составлен авторами по результатам данного исследования

На рис. 9 и 10 изображены первые разности зависимой переменной, предсказанные моделью (4), и истинные первые разности зависимой переменной на валидационных выборках hold-three-out и hold-two-out соответственно.

Результаты применения тестов Зивота – Эндрюса и CIPS оказались несогласованными в одном случае: при проверке исходных рядов на стационарность. Низкая мощность теста CIPS на выборке $T = 14$ не позволяет полагаться на его результаты как на основополагающие; в настоящей работе он использовался в качестве вспомогательного инструмента для подтверждения выводов теста Зивота – Эндрюса. В то же время нельзя полностью исключить возможность стационарности отдельных рядов, однако это не ставит под сомнение общие результаты, поскольку тест Вестерлунда очень твердо подтверждает наличие коинтеграции по обоим релевантным статистикам. Что касается остатков модели (3), оба теста указывают на их нестационарность, что является следствием того, что не все регионы Дальневосточного федерального округа следуют одной модели, из-за чего модель CCEMG, призванная обобщать, подстраивается под динамику аномальных регионов. В связи с этим модель (3) исключается из рассмотрения. Лучше применить или (2), или (1), хотя на графике видно, что (1) вовсе

не способна производить прогнозы, опирающиеся на наблюдаемую динамику.

Долгосрочное уравнение равновесия характеризуется хорошим значением коэффициента детерминации ($R^2 = 0,776$).

Все оцененные коэффициенты при ЕСТ удовлетворяют теоретическим требованиям: они отрицательны и по модулю меньше единицы. Несмотря на адекватность моделей (1) и (2) по hold-out диагностике, появляются вопросы к тому, что коэффициенты коррекции ошибок очень малы. На визуальном анализе рис. 1–3 заметны закономерности, которые портят hold-out диагностику. Модель (1) очень «аккуратно» аппроксимирует динамику изменений, она оказывается лучше, чем наивная, только в силу того, что она почти не двигается (а наивная ошибается на каждом шаге, говоря, что изменение в зависимой переменной равно 0). Модель (2), несмотря на самый низкий из всех R^2 , способна аппроксимировать динамику, она обоснованно лучше, чем наивная модель, на примерно 20%, однако страдает от гетерогенности наклонных коэффициентов. Несмотря на наивысший из всех моделей R^2 , модель (3) очень склонна объяснять данные, которые для общего случая являются «шумом». Среди всей выборки панелей не все регионы следуют заданной коррекции ошибок, и это отражается как на долгосрочном уравнении, так и на модели (3).

Построение модели (4) опирается на наблюдение о наличии регионов-девиантов и выявление этих регионов с помощью алгоритмов кластеризации, для работы которых необходимы подходящие признаки. Результаты кластеризации совпадают с эмпирическими наблюдениями, что, совместно с достаточной дальностью аномальных точек от суперкластера на визуализации с помощью t-SNE, указывает на адекватность выделения аномальных регионов алгоритмом HDBSCAN. Значения метрики silhouette score ниже 0,5 считаются за слабую структуру разделения, что дает основания полагать, что объединение в суперкластер обоснованно, наблюдаемая на рис. 6 относительная близость кластеров друг к другу тоже поддерживает решение о группировке. Как результат аккуратного выделения регионов-девиантов, получена модель (4), которая лидирует по ключевой метрике MASE, показывая себя лучше наивной модели на 45 %.

Коэффициенты при разностях экзогенных переменных интерпретируются как краткосрочные эластичности, ЕСТ интерпретируется как процент девиации от долгосрочного равновесия, который рынок восстанавливает за 1 период. Для наглядности приведем полную интерпретацию для модели (4):

- ЕСТ: за 1 год расхождение между фактическим и долгосрочным равновесным значениями логарифмированного объема предложения сокращается на 35,5 %;

- Δx_{2t} : изменение оборота розничной торговли на 1 % ассоциировано с изменением объема предложения на 0,13 % в тот же период;

- Δx_{1t} : изменение инвестиций в основной капитал на 1 % ассоциировано с изменением объема предложения на 0,293 % в тот же период.

Заключение

В работе предложена эконометрическая модель предложения на рынке жилья Дальневосточного федерального округа на основе панельной модели коррекции ошибок. Применение ЕСМ позволило нивелировать риск ложной регрессии, характерный для стандартных методов оценки при работе с нестационарными и коинтегрированными временными рядами, и разделить краткосрочные эффекты экзогенных переменных от их долгосрочного влияния на эндогенную переменную.

На основе результатов спецификационных тестов предпочтение отдано подходу ССЕМГ в случае подвыборки из восьми регионов, описанных в статье, который

позволяет учитывать гетерогенность коэффициентов и фильтровать влияние общих ненаблюдаемых факторов с помощью кросс-секционных средних. Данная спецификация (модель 4) продемонстрировала наилучшие показатели качества среди четырех оцененных моделей ($R^2 = 0,672$), но применима только к восьми определенным регионам. В данной спецификации присутствуют панельные средние за тот же период – при прогнозировании эта проблема решается одним из двух способов: или строится AR(1) модель для прогнозирования, или, в качестве прокси, подставляются лагированные значения панельных средних. Оба метода работают со сравнимой точностью и корректны с точки зрения теории Песарана и Ямагаты.

На основе визуального анализа, в случае оценки по всем 11 панелям модель (3) подстраивается под аномальные для общей картины регионы, модель (1) лучше наивной только по случайности, пользоваться ей не рекомендуется. Модель (2) оказывается лучшей в случае оценки по 11 панелям, несмотря на $R^2 = 0,206$, запечатляет краткосрочную динамику и способна к предсказанию на основе диагностики по метрикам.

Оценки краткосрочных эластичностей показывают, что в том же периоде изменение оборота розничной торговли на 1 % сопровождается изменением объема предложения жилья на 0,13 % (модель ССЕМГ (4)). Изменение инвестиций в основной капитал на 1 % сопровождается изменением объема предложения жилья на 0,293 % (модель ССЕМГ (4)). Коэффициент при корректирующем члене ЕСТ (–0,355) свидетельствует о том, что в течение одного года рынок устраняет около 35,5 % отклонений от долгосрочного равновесия, вызванных краткосрочными шоками.

Построенное долгосрочное уравнение равновесия со случайными эффектами характеризуется высоким качеством аппроксимации ($R^2 = 0,944$).

Проверка остатков на стационарность подтвердила адекватность моделей МГ (1) и ССЕР (2); модель ССЕМГ (3) исключена из рассмотрения. Модель ССЕМГ (4) рекомендуется к использованию на соответствующих регионах. Модель (2) лучше на оценке по всем панелям. По восьми более гомогенным регионам модель (4) не только лучше по метрикам, но и при увеличении обучающей выборки улучшает метрики качества, в отличие от остальных моделей.

Таким образом, представленные модели могут быть использованы для анализа и прогнозирования предложения на рынке жилья Дальневосточного федерального

округа, в случае модели 4 – строго в рамках оставшихся восьми регионов. Дальнейшие исследования могут быть направлены на расширение набора факторов, учет пространственных эффектов, применение альтернативных подходов к кластеризации регионов и внедрение методов нивелирования зашумления панельных средних, возникающего из-за изоляции подвыборки от регионов-девиантов, применение альтернативных методов оценивания долгосрочных коэффициентов.

Список литературы

1. Сергеева А. О., Измайлова Е. В. Прогнозирование цен на рынке первичного жилья в Санкт-Петербурге на основе регрессионно-дифференциального моделирования // Прикладная математика и вопросы управления. 2019. № 4. С. 157–167. URL: [elibrary_41863873_95146462.pdf](#) (дата обращения: 24.06.2026). DOI: 10.15593/2499-9873/2019.4.10.
2. Токарев Ю. А. Экспертно-статистическое прогнозирование рынка жилья в Российской Федерации // Гуманитарный научный журнал. 2020. № 1. С. 119–126. URL: [elibrary_42813751_10002419.pdf](#) (дата обращения: 24.06.2026).
3. Боченина М. В. Прогнозирование цен на рынке жилья в условиях изменения основной тенденции // Теория и практика общественного развития. 2023. № 8 (184). С. 137–142. URL: [elibrary_54380905_49304635.pdf](#) (дата обращения: 24.06.2026). DOI: 10.24158/tipor.2023.8.16.
4. Харламов А. В., Захаров А. В. Временной анализ эконометрических моделей ценообразования на рынке жилья // Имущественные отношения в Российской Федерации. 2019. № 3 (210). С. 41–50. URL: [elibrary_37136325_96632363.pdf](#) (дата обращения: 24.06.2026). DOI: 10.24411/2072-4098-2019-10301.
5. Филонова Е. С., Жукова А. А. Эконометрический анализ и прогнозирование рынка жилья в муниципальном образовании «Город Орел» // Вестник ОрелГИЭТ. 2018. № 1 (43). С. 84–95. URL: [elibrary_35033284_28144967.pdf](#) (дата обращения: 24.06.2026).
6. Сомов В. Л., Баканач О. В., Цыпин А. П. Эконометрическое моделирование закономерностей рынка арендного жилья в Приволжском федеральном округе // Вестник Саратовского государственного социально-экономического университета. 2018. № 5 (74). С. 200–204. URL: [elibrary_36769769_59127786.pdf](#) (дата обращения: 24.06.2026).

7. Куровский С. В., Мишин Д. А., Амелин Б. М. Эконометрическое моделирование факторов, влияющих на региональные цены на жилье // Мягкие измерения и вычисления. 2025. Т. 93. № 8. С. 114–128. URL: [elibrary_82854210_10253732.pdf](#) (дата обращения: 24.06.2026). DOI: 10.36871/26189976.2025.08.010.
8. Сенникова А. Е., Старцева А. А., Федоров Н. Б., Шаман А. С. Эконометрическое моделирование спроса на жилье в условиях изменения процентных ставок // Вестник Академии знаний. 2025. № 2 (67). С. 537–540. URL: [elibrary_82439276_28352334.pdf](#) (дата обращения: 24.06.2026).
9. Котлярова С. Н., Матушкина Н. А., Мариев О. С. Эконометрическое моделирование нелинейного влияния ипотечного кредитования на строительство жилья в российских регионах // AlterEconomics. 2024. Т. 21. № 4. С. 883–900. URL: [elibrary_77515669_67909982.pdf](#) (дата обращения: 24.06.2026). DOI: 10.31063/AlterEconomics/2024.21-4.12.
10. Мерзлякова А. Ю., Коршунова Е. Ю. Математическая модель первичного рынка жилья города Ижевска // Интеллектуальные системы в производстве. 2017. Т. 15. № 2. С. 4–9. URL: [elibrary_29783633_13277718.pdf](#) (дата обращения: 24.06.2026). DOI: 10.22213/2410-9304-2017-2-4-9.
11. Шабашев В. А., Шорохов С. И., Верхожина М. Ф., Челомбитко А. Н. Экономические, социальные и демографические факторы миграционной привлекательности российских регионов // Региональная экономика: теория и практика. 2017. Т. 15. № 2 (437). С. 391–404. URL: [elibrary_28289681_68758675.pdf](#) (дата обращения: 24.06.2026). DOI: 10.24891/re.15.2.391.
12. Alexander Chudik, M. Hashem Pesaran. Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors // Journal of Econometrics, October 2015. Vol. 188. Is. 2. P. 393–420. URL: [cesifo_wp4232.pdf](#) – Yandex Documents (дата обращения: 24.06.2026). DOI: 10.1016/j.jeconom.2015.03.007.
13. Hashem M. Pesaran. A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence // Journal of Applied Econometrics, March 2007. Vol. 22. Is. 2. P. 265–312. URL: <https://www.jstor.org/stable/25146517> (дата обращения: 24.06.2026). DOI: 10.1002/jae.951.
14. Joakim Westerlund. Testing for cointegration in panel data. // Oxford Bulletin of Economics and Statistics, December 2007. Vol. 69. Is. 6. P. 709–748. DOI: 10.1111/j.1468-0084.2007.00477.x.
15. Sean Holly, M. Hashem Pesaran, Takashi Yamagata. A Spatio-Temporal Model of House Prices in the US // Journal of Econometrics, September 2010, Vol. 158. Is. 1. P. 160–173. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0304407610000837> (дата обращения: 14.01.2026). DOI: 10.1016/j.jeconom.2010.03.040.

Конфликт интересов: Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest: The authors declare that there is no conflict of interest.

Финансирование: Авторы заявляют об отсутствии внешнего финансирования.

Financing: The research was performed without external funding.