ИКОНОМЕТРИЧЕН МОДЕЛ НА ЗАВИСИМОСТТА НА НОВОПОСТРОЕНИТЕ ЖИЛИЩА ОТ ОСНОВНИ ФАКТОРИ: АНАЛИЗ ЧРЕЗ ПАНЕЛНИ ДАННИ ЗА БЪЛГАРСКИТЕ РЕГИОНИ

Весела Димитрова

ECONOMETRIC MODEL OF THE NEWLY BUILT DWELLINGS AS A FUNCTION OF MAJOR FACTORS: ANALYSIS OF PANEL DATA FOR BULGARIAN REGIONS

Vesela Dimitrova

Abstract: *The article suggests selected alternative models of the dependence of newly built dwellings on selected socio-economic factors evaluated by regions in Bulgaria. The study period is limited to 2005 – 2021 and covers periods of economic growth, financial and economic crisis, recovery, and subsequent global crisis caused by the COVID-19 pandemic. Alternative specifications of econometric models are constructed and estimated using panel data for Bulgarian districts regarding the living area of newly built dwellings together with selected demographic, social and economic factors. A comparative analysis of the results from models estimated with different panel specifications – „pooled OLS method", „fixed-effects method" and „random effects method" – is proposed.*

Key words: new dwellings; panel data regression; Bulgarian regions.

Увод

През последните две десетилетия строителството на нови жилища в България се развива доста динамично. В строителния сектор се реализира значителна по обем продукция за периода на икономически растеж от 2005 – 2008 г., както за инженерното (ниско) строителство, така и по отношение на сградното (високо) строителство. В резултат на настъпилата финансова и икономическа криза през 2008 г., негативната тенденция в жилищното строителство се запазва за периода 2009 – 2015 г. Ръстът на жилищното строителството през последните години е значителен, въпреки инфлацията, пандемията от COVID-19 и войната в Украйна. България попада в челните места по брой на новопостроени жилища на човек от населението и по брой на новопостроени жилища сред европейските страни.

Започналите строително-монтажни дейности и издадените разрешителни за строеж за нови сгради се увеличават също за периода 2018 – 2022 г., което може да доведе до още по-голям бум на жилищното строителство през следващите няколко години. Този бум в жилищния сектор може да предизвика следваща криза в строителството, като тази през 2008 г., в резултат на което много жилища и сгради могат да останат празни. От особена важност е да се изследва зависимостта между динамиката на новопостроените жилища и нейни основни демографски, социални и икономически фактори.

Обект на изследването е полезната жилищна площ в новопостроените жилища по области в България за периода 2005 – 2021 г. включително. Използвани са официални годишни статистически данни за жилищния фонд и регионалните демографски, социални и икономически показатели, публикувани в уебсайта на НСИ.

Основната цел на статията е да се изследва зависимостта на новопостроените жилища в българските региони от очакваното влияние на избрани демографски и социално-икономически фактори. За постигането на целта в изследването са очертани следните конкретни задачи:

– Систематизиране на фактори с очаквано влияние върху полезната жилищна площ в новопостроените жилища на регионално ниво;

– Съставяне на информационен масив от панелни данни за полезната жилищна площ в новопостроените жилища в българските региони и за факторите с очаквано влияние за периода 2005 – 2021 г.;

– Построяване на иконометрични модели при панелни данни за полезната жилищна площ в новопостроените жилища;

– Оценяване на иконометрични панелни модели на зависимостта на полезната жилищна площ от очакваните фактори;

– Интерпретация на получените емпирични резултати от тези модели.

1. Основни демографски и социално-икономически фактори, влияещи върху новопостроените жилища

Полезната жилищна площ е измерителят, с който се характеризира най-адекватно продукцията в жилищното строителство. Според методологията на НСИ, полезната площ в новопостроените жилища представлява сума от жилищната и спомагателната площ. Жилищната площ, от своя страна, включва площта на стаите за живеене, спалните, нишите за спане, трапезарите, стаите за дневен престой, стаите ползвани като работни кабинети и библиотеки, гостните, холовете, както и площта на кухните над 4 кв. м. В спомагателната площ се включва площта на спомагателните помещения, стаите и кухните с площ, по-малка от 4 кв. м, вестибюлите с портал и друга преграда, коридорите, антретата, баните, тоалетните, килерите, дрешниците, другите спомагателни помещения (сушилни, перални, балкони и лоджии) независимо от големината на площта им[1].

Върху жилищното строителство оказват влияние редица фактори. Няколко разработки изследват факторните променливи, влияещи върху тенденцията на пазара за недвижими имоти за отделни региони в страната (Атанасова, 2021; Томова, 2018). Атанасова (2021) изследва демографските промени върху цените и обема на строителната продукция в област Варна. Според Томова (2018) факторите, влияещи върху пазара на недвижими имоти, могат да бъдат разделени на общи и специфични. Общите фактори се класифицират като демографски, икономически, политически, технологични, природни, културни и други (Томова, 2018).

За целите на изследването ще разделим факторите, влияещи върху новопостроените жилищни площи, на две основни групи: демографски и социално-икономически. Политическите, технологичните, културните и други фактори са извън обхвата на изследването в тази разработка.

Демографските фактори, които разглеждаме, са показателя за естествен прираст и показателя за механичното движение на населението. Естественият прираст на населението представлява разликата между броя на регистрираните живородени деца и броя на умрелите през годината. Коефициентът за естествен прираст показва броя на увеличението (намалението) на населението на 1 000 души, като отношение на разликата между броя на живородените и броя на умрелите лица през годината и средногодишния брой на населението през същата година[2]. Това е потенциалното население през следващите години, което ще живее, учи и работи в съответния район или област. Когато коефициентът на естествен прираст е положителен, това показва, че броят на ражданията е по-голям от броя на смъртните случай. За България този показател е отрицателен за целия период на изследването 2005 – 2021 г., особено през двете години на пандемията от COVID-19, когато броят на смъртните случай нарасна значително. Най-засегнати са областите Велико Търново, Видин, Враца, Габрово, Кюстендил, Ловеч, Перник, Търговище, Ямбол. Това са области, в които не достигаха лекари, медицински персонал и болнични легла по време на пандемията от COVID-19.

При изследване на миграцията се наблюдава съвкупността на лицата, променили обичайното си местоживеене. Показателят за механичния прираст на населението в териториален аспект представлява разликата между броя на заселените и броя на изселените лица във и от всяко населено място[3]. Коефициентът за механичен прираст показва броя на увеличението (намалението) на населението на 1 хил. души, като отношение на разликата между броя на заселените и броя на изселените лица през годината и средногодишния брой на населението през същата година. Механичният прираст на населението е другият демографски фактор, който би могъл да оказва влияние върху размера на новопостроените жилища. Колкото повече се увеличава населението в резултат на заселилите се лица, толкова повече се очаква да бъде търсенето на жилища, в това число и на новопостроените жилища в съответния район.

Заселването на хората в по-големите градове и в икономически по-развитите райони на България е тенденция, която се наблюдава през последните няколко десетилетия. Изключение, може би са годините с пандемията от COVID-19, когато част от населението търси жилища в по-малките населени места, за да има свобода от прилаганите мерки при временните ограничения от пандемията.

От демографските фактори, които могат да бъдат включени при изследване на зависимостта на новопостроената жилищна площ, е частта на населението във възрастовата група над 18 години, като потенциално население, което би могло да закупи новопостроени жилища.

Основните социални фактори, определящи търсенето на жилища в даден регион, са здравеопазването, образованието, пътната инфраструктура и други.

Здравеопазването може би е един от най-основните социални фактори, които могат да се включат при анализа на новопостроените жилища. Показателите, чрез които се характеризира статистическото наблюдение на общественото здраве в България са лечебните заведения за болнична и извънболнична помощ, здравните заведения, съвкупността на медицинския персонал, легловата база и др. Здравните и лечебните заведения не са свързани пряко с обслужването на населението. От една страна по-важен е показателят за съвкупността от медицински персонал, който обслужва даден регион, отколкото броят на здравните и лечебни заведения. От друга страна голяма част от лечебните и здравните заведения обслужват една област, но специализираните заведения обслужват населението на две или три области. Легловият фонд е друг съществен показател за здравеопазването, но без лекари и здравен персонал в лечебните заведения не би могла да се осъществява здравната грижа. Пандемията от COVID-19 бе доста показателна за недостига на лекари, здравен персонал, болнични легла в здравните и лечебните заведения. Осигуреността на даден регион с лекари в лечебните и здравни заведения, като фактор, би могла да доведе до ръст в новопостроените жилища.

При предлагането и търсенето на жилища един от основните критерии е осигуреността на съответния регион с образователни институции. Образователните институции, за които се събират статистически данни, са детските градини, училищата за начално, прогимназиално и гимназиално образование, професионалните колежи и висшите училища.

Образованието позволява на хората да развиват умения си, което води до увеличаване на заетостта и икономическия растеж. Положителните прирасти на БВП и доходите на икономически активните единици в икономиката се отразява благоприятно върху свободните финансови ресурси, които от своя страна могат да увеличат банковите депозити, които могат да бъдат инвестирани в жилищни имоти или във финансови инструменти.

Пътищата осигуряват връзките между населените места. Особено важни са както дължината на пътната мрежа, така и вида и качеството на самата настилка. Статистическият показател, който характеризира пътната инфраструктура, е дължина на пътищата, като на ниво общини и области се включва дължината на първокласните, второкласните и третокласните пътища.

Строителството на нови жилища се влияе и от редица икономически фактори. Те са свързани с икономическата активност в страната и по региони, заетостта, безработицата, средната цена на предлагане на новопостроените жилища, доходите и спестяванията на населението и др.

Както търсенето, така и предлагането на новопостроени жилища за българските региони се очаква да нараства в резултат на повишената икономическа активност. Измерването на икономическата активност би могло да се оцени с равнището на БВП. Най-показателен оценител на ниво области и общини е добавената стойност по факторни разходи. Това е стойността на произведената продукция, от която е приспадната стойността на акцизите и разходите за оперативната дейност, без тези за персонала, и са добавени приходите от финансирания. Използвайки данните за наетите лица по трудови и служебни правоотношения, е възможно и включването на производителността на труда при анализа на взаимодействието на новопостроените жилища и икономическите фактори, които им влияят.

2. Методи на изследването

Развитието на иконометричния софтуер през последните години доведе до широко приложение на методите за анализ на панелни данни в различни области на икономиката, социалната сфера, социологията и др. Предимствата на панелните изследвания са, че носят висока информативност за изследваните процеси и чрез тях се контролира т.нар. „индивидуална хетерогенност“ или още „разнородност“ (Baltagi, 2005; Hsiao, 2003). Чрез панелните модели е възможно по-детайлно изследване на единиците и динамиката на тяхното поведение, както и използване на по сложни иконометрични модели за идентификация на икономически ефекти (Hsiao, 2003).

Общият регресионен модел на панелните данни може да бъде изразен по следния начин:

$$Y_{it} = \beta_0 + \sum_{j=1}^{k} \beta_j X_{jit} + v_{it} \quad (1)$$

където:
i – индекс за включените променливи;
t – индекс за времевия период на измерване;
k – брой включени обяснителни променливи;
β_0 – константа (свободен член) на модела;
β_j – коефициент на променлива "j" (j=1…k);
Y_i – зависимата променлива;
X_j – включена обясняваща променлива „j";
v_{it} – грешка на модела.

Грешката (остатъчен компонент) v_{it} в уравнение (1) може да бъде разложена на две:

$$v_{it} = w_{it} + \omega_i$$

където ω_i е компонент на грешката, специфичен за панелната единица "i". Тази грешка не се променя с течение на времето (Baltagi, 2005; Greene, 2002; Gujarati, 2009; Wooldridge, 2003). В литературата се нарича още ненаблюдаван ефект, ненаблюдавана „хетерогенност", „фиксирани ефекти" (cross-sectional unit-specific error). Тези ефекти са специфични (индивидуални) при отделните единици от наблюдаваната съвкупност (поради специфика на локацията, наличните ресурси, сравнителни предимства/недостатъци и др.).

Компонентът w_{it} представлява ненаблюдавана, но вариираща с времето грешка. Тя участва в стойността на зависимата променлива, тъй като обхваща влиянието на невключените в модела фактори, които варират с течение на времето, например свързани с промени в икономическите условия (idiosyncratic error).

Ползата от разлагането на общата грешка на две части е, че ако може да се елиминира част от тях с помощта на панелните данни, ще могат да се редуцират грешките, причинени от пропускането на невключени фактори, които са специфични за панелните единици.

Показано е, че регресионните модели ще имат неефективни оценки на коефициентите β_j, ако ненаблюдаемите ефекти ω_i не се отстранят. Когато те са корелирани с някоя от независимите променливи X_j, се наблюдава изместване на тези оценки, предизвикано от съществуващата хетерогенност (нееднородност) на панелните единици. Ако независимите променливите X_j обхващат всички важни характеристики на единиците, тогава ω_i няма да варират съществено и следователно може да се приложи модел на „обединени данни в статика и динамика", при което като ω_i ще бъдат изключени от регресията (Маринов, 2014).

Един от основните и лесни методи за оценка на уравнение (1) е простото обединяване на данните и прилагане на метода на най-малките квадрати (pooled OLS). За да се оцени уравнение (1), използвайки този обикновен МНМК, е необходимо да се приеме, че съставната грешка v_{it} не е свързана с обяснителните променливи X_{jit} (Greene, 2002; Gujarati, 2009; Wooldridge, 2003). При „обединения модел" всички двойки „единица-период" [i,t] се разглеждат като независими наблюдения от една и съща извадка. Този модел е известен още като „модел с константни коефициенти" (constant coefficient model). Това означава, че само когато няма нито индивидуални, нито времеви фиксирани ефекти, можем да оценим регресионни модели с „обединени данни в статика и динамика".

Горепосочената спецификация има някои недостатъци. Този метод пренебрегва детайлната информация, носена от панелните данни. Допускането тук често не е реалистично, тъй като не е възможно да се обхванат и включат в модела всички променливи, обясняващи постоянните във времето и специфични за единица "i" ефекти ω_i. Поради тази традиционна причина използването на „обединения модел" при анализиране на панелни данни не се счита за оправдано (Gujarati, 2009; Wooldridge, 2003).

„Моделът с фиксираните ефекти" (fixed effects model) се използва широко, когато искаме да контролираме пропуснатите променливи, които имат постоянно действие през периода от време и варират само между панелните единици. Тези компоненти се определят като „фиксирани ефекти" и отразяват съществуващата ненаблюдавана хетерогенност. Общата спецификация на този модел е представена чрез уравнение (2).

$$Y_{it} = \beta_0 + \sum_{j=1}^{k} \beta_j X_{jit} + \mu_t + \omega_i + u_{it} \qquad (2)$$

При оценяване на уравнение (2) се приема, че компонентите на ненаблюдаваната хетерогенност (ω_i) корелират с една или повече от обяснителните променливи (X_{jit}). Друго важно предположение е, че идиосинкратичната грешка (w_t) е независима от обясняващите променливи (X_{jit}) (Baltagi, 2005; Wooldridge, 2003). Чрез елиминиране на ненаблюдаваните ефекти ω_i, което

предполага намаляване на грешките поради пропуснати обяснителни променливи, можем да имаме по-ефективни оценки на параметрите (т.е. с по-малки стохастични грешки). Аналогична полза би била налице и при елиминиране на т.нар. „темпорални фиксирани ефекти“ μ_t, които отчитат постоянни нива на „Y“ за даден период (t), еднакви за всички панелни единици, но променящи се с течение на времето (t=1,…,T).

В специализираната литература са описани три метода за елиминиране на ненаблюдаемите „фиксирани ефекти“ ω_i и μ_t. В настоящата разработка се прилага най-популярният метод, а именно техниката с включване на фиктивни (дъми) променливи. При нея се работи на принципа на включването на такива променливи както за панелните единици, така и за периодите от време. Ненаблюдаемите ефекти се оценяват като коефициенти пред тези дъми променливи. Може да се получат оценки за ω_i и μ_t, като се включи дъми променлива за всяка единица на напречното сечение (i) и за всеки период от време (t) (Baltagi, 2005; Greene, 2002; Gujarati, 2009). Този метод е известен още като „регресия с фиктивни променлива по метода на наймалките квадрати“ (LSDV: Least-squares Dummy Variables Method).

Ако при използване на панелни данни се заложи допускането, че компонентите ω_i не корелират с обясняващите променливи X_{jit}, то тогава могат да се използват и т.нар. „модели със случайни ефекти“ (REM: Random Effects Models) (Wooldridge, 2003). Според Gujarati (2009) регресионното уравнение тук има следния вид:

$$Y_{it} = \alpha_i + \beta X_{it} + a_i + \varepsilon_{it} \quad (3)$$

където:

α_i – случайна променлива за индивидуалния ефект при единица “i”, със средна стойност α;

ε_{it} – случайна грешка с комбиниран характер (свързана както с единиците, така и с времевите периоди).

Моделът на случайните ефекти се оценява чрез т.нар. Обобщен метод на най-малките квадрати (Generalized LS) или негова модификация, известна като „приложим ОМНМК“ (Feasible GLS) (Greene, 2002). Надеждните резултати от тези по-сложни методи изискват сравнително дълги периоди, както и по-голям брой панелни единици. Тогава ОМНМК дава възможност и за компенсиране на различни трудности, свързани с поведението на остатъците ε_{it} (тези въпроси са извън обхвата на изложението тук).

Проверките за значимостта на фиксираните ефекти в регресионните панелни модели спрямо моделите оценени чрез обединения МНМК (pooled OLS), се базират на теста на Фишер и теста на Валд. Нулевата хипотеза е, че включването на фиктивните променливи не е оправдано; техните коефициенти не варират, или са статистически незначими. При отхвърляне на нулевата хипотеза за едната или другата група от коефициенти за дъми променливи става основателно да се използва моделът на фиксираните ефекти.

Чрез LM-теста на Breusch-Pagan може да се провери хипотезата за значимостта на модел със случайни ефекти спрямо модели, оценени чрез обединения МНМК. Нулевата хипотеза при този тест е, че дисперсията на остатъчните компоненти е нула. При отхвърляне на нулевата хипотеза може да се приеме, че моделът със случайните ефекти е за предпочитане и отчита по-добре проблемите с хетерогенността на данните (Greene, 2002).

Ако диагностичните тестове потвърждават предимствата на моделите с фиксирани и случайни ефекти, окончателният избор се извършва след прилагане на теста на Хаусман (Hausman test). При този тест нулевата хипотеза предполага, че между оценките на параметрите, получени по двете панелни модификации, няма съществена разлика. Тогава моделът със случайни ефекти се приема за по-подходящ „оценител“ поради по-ниските стохастични грешки. Ако H_0 хипотеза бъде отхвърлена, то между двата набора от оценки има съществени различия. В този случай следва, че има корелация между индивидуалните ефекти и поне един от включените в модела предиктори, поради което моделът с фиксирани ефекти става по-подходящ.

3. Информационна осигуреност на масива от панелни данни

Данните в настоящата статия са базирани на информационната статистическа система на НСИ на България. Статистическите данни за полезната жилищна площ в новопостроените жилища, демографските и социално-икономическите фактори са годишни за административните региони (области) на ниво NUTS-3. Основните променливи са

взети директно от платформата с данни на НСИ, а други са трансформирани за нуждите на настоящето изследване. Времевият обхват на изследването е периода 2005 – 2021 г., а пространственият са 27-те области на България, Включени са всички области, с изключение на София (столица). Този регион се отличава значително от останалите области. За столицата се наблюдават съществени различия както в динамиката на новопостроените жилища, така и при всички икономически, демографски и социални показатели. Тук е съсредоточена голяма част от населението и добавената стойност за страната е най-голяма, в сравнение с всички области на България.

В изследването включваме полезната жилищна площ като най-адекватен индикатор за областния размер на новопостроените жилища, за разлика от броя на новопостроените сгради, брой жилища, брой стаи и други в новопостроените имоти.

От демографските фактори в иконометричните модели са включени коефициентът на естествен прираст на населението и коефициентът на механичен прираст. Населението над 18 години не е включено в иконометричните панелни модели, тъй като има висока колинеарност с включените по-горе демографски фактори.

В информационния масив от панелни данни социалните фактори, които се включват, са лекарите в лечебните и здравните заведения, броят на училищата в начален, прогимназиален и гимназиален етап и дължината на първокласните, второкласните и третокласните пътища.

В изследването не включваме легловата база в лечебните и здравните заведения и броя на учениците в училищата поради факта, че има висока мултиколинераност между училищата, броят на учениците, броят на леглата в лечебните заведения и дължината на пътната мрежа.

От икономическите показатели в информационния панелен масив, факторът който отразява адекватно икономическата ефективност, е областната производителност на труда. За получаването на този показател са използвани данни за добавена стойност по факторни разходи и средно списъчния брой на наетите лица по трудово и служебно правоотношение в областите. Тъй като това е трансформиран показател, в който участват наетите лица, факторите за безработица и заетост не се включват в иконометричните панелни модели, поради колинеарността между тях.

В разработката не са разгледани цените на новопостроените жилища, като част от икономическите фактори оказващи влияние върху обема на строителството. За цените на жилищата липсват официални и надеждни данни по отделните региони за целия изследван период 2008 – 2021 г. НСИ събира и публикува данни от 2015 г за цените на жилища и изчислява индекси на пазарните цени на жилища (новопостроени и съществуващи). Данните за периода преди 2014 г. са за „Средни пазарни цени на жилища“.

Използваните означения за променливите в иконометричните модели са:

all_living_space – полезна жилищна площ в новопостроените жилища по области в кв. м. Понататък в разработката използваме понятието „жилищна площ“ като еквивалент на полезната жилищна площ в новопостроените жилища;

rate_of_natur_incr – коефициент на естествен прираст (на 1000 души от населеното);

rate_of_migration_incr – коефициент на механичен прираст на населението (на 1000 души от населеното). От 2007 г. данните за механичния прираст по области са резултат, както от вътрешната, така и от външната миграция;

doctors – лекари в лечебните и здравните заведения на 10 000 души от населеното;

schools – брой училища (начален, прогимназиален и гимназиален етап);

roads – дължина на пътищата в километри;

labour_prod – производителност на труда на едно наето лице по трудово и служебно правоотношение в хил. лв.

Съставеният информационен масив за българските региони относно жилищната площ в новопостроените жилища и променливите, които са очаквани предктори, е за периода 2005 – 2021 г. и представляват балансиран панел от данни.

4. Емпирични резултати от оценените иконометрични модели

4.1. Иконометричен панелен модел оценен с обединения метод на най-малките квадрати (pooled OLS)

Регресионното уравнение на зависимостта между жилищната площ на новопостроените жилища и изследваните предикторни променливи е оценено чрез обединения МНМК (pooled OLS). При направените статистически проверки се установи, че панелният иконометричен модел оценен с pooled OLS е адекватен [Fstat=86.7 и Prob(F)<0,01].

Всички коефициенти пред факторните променливи са статистически значими, като равнищата им на значимост са по-малки от 1% за всички коефициенти (табл. 1). Заключението, което може да се направи е, че 54% от вариацията на жилищната площ на новопостроените жилища може да се обясни чрез оцененият модел на панелна регресия между жилищната площ и изследваните фактори.

$$ln_all_living_space_{it} = \delta_0 + \delta_1 . rate_of_natur_incr_{it} + \delta_2 . rate_of_migration_incr_{it} + \delta_3 . doctors_{it} + \delta_4 . schools_{it} + \delta_5 . roads_{it} + \delta_6 . ln_labour_prod_{it} \quad (4)$$

където: δ_0, δ_1, δ_2, δ_3, δ_4, δ_5, δ_6 – параметри на модела; ***i*** – номер на област; ***t*** – номер на година.

Таблица 1.

Резултатите от многофакторния панелен модел, оценен с обединения МНМК (pooled OLS)

Променливи	Коефициенти	Станд. грешка	t-статистка	Prob.
rate_of_natur_incr	0.1007	0.0131	7.6637	0.0000
rate_of_migration_incr	0.0473	0.0078	6.0539	0.0000
doctors	0.0191	0.0072	2.6658	0.0080
schools	0.0076	0.0018	4.2914	0.0000
roads	0.0010	0.0002	4.1230	0.0000
ln_labour_prod	0.5823	0.1067	5.4564	0.0000
c	6.8551	0.3031	22.6154	0.0000
R-squared	0.5351	F-statistic		86.692
Adjusted R-squared	0.5289	Prob(F-statistic)		0.0000

Оцененият иконометричен панелен модел с обединения МНМК е следния:

$$ln_all_living_space_{it} = 6.8551 + 0.1007.rate_of_natur_incr_{it} + 0.0473.rate_of_migation_incr_{it} + 0.0191.doctors_{it} + 0.0076.schools_{it} + 00010.roads_{it} + 0.5823.ln_labour_prod_{it}$$

Нарастването на коефициента на естествен прираст, миграцията, броят на лекарите в лечебните заведения, броят на училищата, дължината на пътищата и производителността на труда водят до нарастването на новопостроената жилищна площ в областите в България. Трябва да бъдем внимателни преди да приемем този резултат, защото той предполага, че ефектът на даден фактор при всички области е един и същ, измерен като средна величина чрез неговия коефициент. Също така, при всички области е едно и също влиянието на трендовия фактор през целия изследван период.

4.2. Иконометричен панелен модел с фиксирани ефекти (Fixed-Effect Model)

При моделите с фиксираните ефекти е възможно да се включат фиктивни променливи за изследваните периоди и фиктивни променливи, които представят областите.

$$ln_all_living_space_{it} = \delta_0 + \delta_1.rate_of_natur_incr_{it} + \delta_2.rate_of_migration_incr_{it} + \delta_3.doctors_{it} + \delta_4.schools_{it} + \delta_5.roads_{it} + \delta_6.ln_labour_prod_{it} + \sum_{m=2}^{17} \gamma_m DT_{mit} + \sum_{r=2}^{27} \beta_r DR_{rit} \quad (5)$$

където:

γ_m – параметрите измерват фиксираните ефекти за периодите от 2006 – 2021 г., без първата година, която е избрана за референтна;

β_r – параметрите измерват фиксираните ефекти за отделните области, без първата, която е избрана за референтна.

От експериментираните иконометрични модели с фиксирани ефекти се оказа, че най- удачни са моделите с включването на фиктивни променливи само относно изследваните периоди (табл.2).

Таблица 2.

Резултатите от многофакторния панелен модел, оценен с фиксирани ефекти (FEM)

Променливи	Коефициенти	Станд. грешка	t-статистка	Prob.
rate_of_natur_incr	0.1462	0.0165	8.8758	0.0000
rate_of_migration_incr	0.0467	0.0083	5.5915	0.0000
doctors	0.0316	0.0072	4.3950	0.0000
schools	0.0040	0.0019	2.1297	0.0338
roads	0.0012	0.0002	5.1302	0.0000
ln_labour_prod	0.4322	0.1704	2.5369	0.0115
DT_2	……	……	……	……...
…….	……	……	……	……...
DT_17	……	……	……	……...
c	7.3522	0.4498	16.3463	0.0000
Period fixed (dummy variables)				
R-squared	0.6128	F-statistic		31.370
Adjusted R-squared	0.5933	Prob(F-statistic)		0.0000

Резултатите показват, че обяснителната способност на иконометричния модел с фиксираните ефекти се увеличава спрямо иконометричният модел, оценен чрез обединения метод на най-малките квадрати (pooled OLS) на 61%. При статистическата проверка се установи, че панелният иконометричен модел с фиксираните ефекти *(FEM)* е адекватен [*Fstat*=31.4 и Prob(F)<0,01]. Всички коефициенти пред факторните променливи са статистически значими, като равнищата им на значимост са по-малки от 1% за всички коефициенти, само за коефициента пред променливата „брой училища“ равнището е по-малко от 5%.

Оценениятиконометричен панелен модел с фиксирани ефекти е следния:

$$ln_all_living_space_{it} = 7.3522 + 0.1462.rate_of_natur_incr_{it} + \\ + 0.0467.rate_of_migration_incr_{it} + 0.0316.doctors_{it} + 0.0040.schools_{it} \\ +0.0012.roads_{it} + 0.4322.ln_labour_prod_{it} + \sum_{m=2}^{17} \gamma_m DT_{mit}$$

За да изберем кой от двата иконометрични модел да използваме при оценяване на зависимостта между жилищната площ в новопостроените жилища и изследваните фактори ще проверим съвместната значимост на годишните ефекти чрез тест на Валд (Wald), като проверката се извършва спрямо параметрите, които са оценени с обединения метод на най-малките квадрати (pooled OLS). Нулевата хипотеза е, че параметрите за всички фиктивни променливи, с изключение на една, са нула т.е. H_0: $\gamma_2 = \ldots = \gamma_m = 0$. При направените статистически проверки при теста на Валд (Wald) се установи, че отхвърляме нулевата хипотеза [χ^2=87.6 и Prob(χ^2)<0.01], следователно моделът с включени фиксирани ефекти за годините може да се приеме за по-подходящ.

4.3. Многофакторен иконометричен панелен модел, оценен със случайните ефекти (Random Effects Model – REM)

При моделите със случайни ефекти е характерно, че влиянието на отделните области и изследваните периоди са включени в неявен вид, и че тези фиксирани ефекти са независими и не корелират с факторните променливи включени в иконометричния модел.

$$ln_all_living_space_{it} = \delta_0 + \delta_1.rate_of_natur_incr_{it} + \delta_2.rate_of_migration_incr_{it} + \\ + \delta_3 .\ doctors_{it} + \delta_4.schools_{it} + \delta_5.roads_{it} + \delta_6.ln_labour_prod_{it} + \ \varepsilon_it \quad (6)$$

Оценките на параметрите на иконометричния модел със случайните ефекти са получени чрез обобщения метод на най-малките квадрати (Generalized least squares – GLS). Този метод постига и справяне с хетероскедатичността, ако такава съществува при остатъчния компонент.

Таблица 3.

Резултатите от многофакторния панелен модел, оценен със случайните ефекти (REM)

Променливи	Коефициенти	Станд. грешка	t-статистка	Prob.
rate_of_natur_incr	0.1112	0.0132	8.4127	0.0000
rate_of_migration_incr	0.0481	0.0076	6.3496	0.0000
doctors	0.0223	0.0068	3.2834	0.0011
schools	0.0067	0.0017	3.9126	0.0001
roads	0.0011	0.0002	4.5543	0.0000
ln_labour_prod	0.5749	0.1106	5.1973	0.0000
c	6.9053	0.3091	22.3417	0.0000
R-squared	0.5532	F-statistic		93.269
Adjusted R-squared	0.5473	Prob(F-statistic)		0.0000

Коефициентът на детерминация на иконометричния модел със случайните ефекти е 55 %, който е по-нисък спрямо иконометричният модел, оценен чрез фиксираните ефекти. Панелният модел със случайните ефекти *(REM)* е адекватен [*Fstat*=93.3 и Prob(F)<0,01]. Всички коефициенти пред факторните променливи са статистически значими, като равнищата им на значимост са по-малки от 1% за всички коефициенти.

Оцененият иконометричен панелен модел със случайните ефекти е следния:

$$ln_all_living_space_{it} = 6.9053 + 0.1112.rate_of_natur_incr_{it} + \\ + 0.0481.rate_of_migratio_incr_{it} + 0.0223.doctors_{it} + 0.0067.schools_{it} + 0.0011.roads_{it} \\ + 0.5749.ln_labour_prod_{it}$$

Чрез LM-теста на Breusch-Pagan проверяваме хипотезата за значимостта на модели със случайни ефекти спрямо модели, оценени чрез обединения МНМК (pooled OLS). Нулевата хипотеза при този тест е, че дисперсията на остатъчните компоненти е нула. Емпиричната характеристика на BP-LM теста е 1068,56 и при риск за грешка, по-малка от 1% [Prob(BP-LM)<0,01] се отхвърля нулевата хипотеза. Следователно може да се приеме, че моделът със случайните ефекти е за предпочитане пред модела, оценен чрез обединения МНМК; от своя страна, моделът със случайни ефекти отчита по-добре проблемите с хетерогенността на данните

За да изберем иконометричен модел с фиксирани ефекти или такъв със случайни ефекти за зависимостта на новопостроените жилищни площи от изследваните фактори прилагаме теста на Хаусман (Hausman test). Този тест определя дали оценките на коефициентите, взети като група, съществено се различават при двата варианта на панелната регресия. Нулевата хипотеза гласи, че моделът със случайните ефекти (REM) е по-подходящ оценител. Резултатите от този тест са представени в табл. 4.

Таблица 4.

Резултати от приложението на Хаусман теста (Hausman test)

Променливи	Fixed effects коефициенти	Random effects коефициенти	Var(Diff.)	Prob.
rate_of_natur_incr	0.1462	0.1112	0.0001	0.0004
rate_of_migration_incr	0.0467	0.0481	0.0000	0.6887
doctors	0.0316	0.0223	0.0000	0.0001
schools	0.0040	0.0067	0.0000	0.0015
roads	0.0012	0.0011	0.0000	0.0088
ln_labour_prod	0.4322	0.5749	0.0168	0.2710
Hausman test	χ^2 (6) = 47.738			0.0000

Емпиричната характеристика на теста на Hausman е χ^2=47.7 и при равнище на значимост по-малко от 1% [Prob(χ^2)<0,01] се отхвърля нулевата хипотеза, следователно може да се заключи, че по-подходящ е иконометричният модел с фиксираните ефекти (FEM).

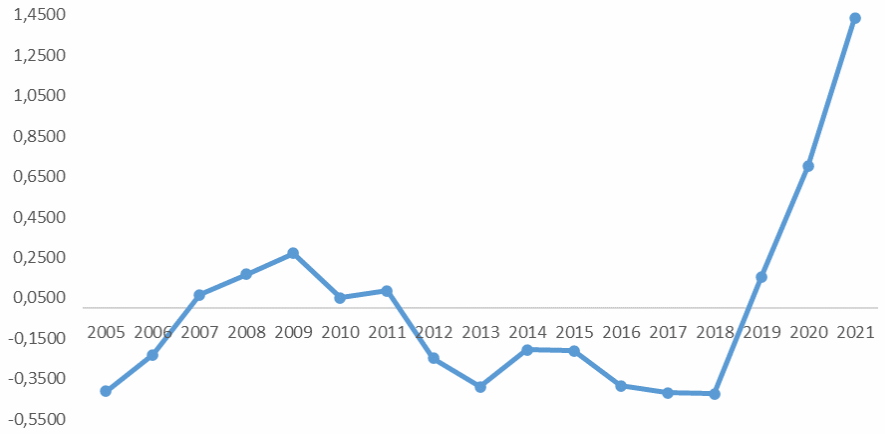
Като вземем под внимание характера на коефициентите на иконометричния панелен модел с фиксираните ефекти, а именно коефициенти на полуеластичност и еластичност, можем да направим следните изводи относно закономерностите за периода 2005 – 2021 г.:

- На 1‰ по високо ниво на естествения прираст, жилищната площ при новопостроените жилища в областите нараства средно с 14.6%.
- На 1‰ по високо ниво на механичния прираст, жилищната площ при новопостроените жилища в областите нараства средно с 4.7%.
- При увеличаване на лекарите в лечебните и здравните заведение (средно на 10 хил. души) с един лекар, жилищната площ при новопостроените жилища в областите нараства средно с 3.2%.
- При увеличаване на училищата в начален, прогимназиален и гимназиален етап с 1 училище, жилищната площ при новопостроените жилища в областите нараства средно с 0.4%.
- При увеличаване на средната производителност на труда на ниво област с 1%, жилищната площ в новопостроените жилища нараства с 0.43%.

Въпреки, че е статистически значим коефициентът пред променливата „дължината на пътната мрежа“, тя оказва незначително влияние върху изменението в новопостроената жилищна площ (0.1%).

Всички тези изводи са валидни при елиминирането на промените в средните равнища на жилищната площ в новопостроените жилища под влиянието на трендовите фактори през изследвания период.

Трендовите ефекти върху новопостроената жилищна площ могат да се видят на фиг.1. Ясно са очертани негативните тенденции в жилищното строителство през годините 2005 и 2006 г. и за периода 2012 – 2018 г.



Фиг. 1. Влиянието на фиксираните времеви ефекти върху жилищната площ в новопостроените жилища по области за периода 2005 – 2021 г.

Заключение

Изследването представя резултатите, получени чрез оценяване на иконометрични модели на зависимостта на жилищната площ в новопостроените жилища по региони в България за периода 2005 – 2021 г. Тук фокусът попада върху основни демографски и социално-икономически фактори, които влияят върху размера на новопостроената жилищна площ.

От оценените иконометрични панелни модели се установи, че спецификацията с фиксираните ефекти е по-добър модел на зависимостта на размера на жилищната площ в новопостроените жилища от естествения прираст, механичния прираст, броя на лекарите в лечебните и здравни заведение, броя на училищата, пътната инфраструктура и производителността на труда. Въпреки това, неговата обяснителна способност е недостатъчно висока, за да разкрие цялостната картина

в изменението на новопостроените жилищни площи. Ефектите на факторите са оценени с коефициентите на полуеластичност и еластичност, като се оказа, че влиянието на дължината на пътната мрежа, оказва незначително влияние върху новопостроената жилищна площ (0.1%). Най-голямо е изменението на жилищната площ в новопостроените жилища под влиянието на факторите „естествен прираст", „механичен прираст" и „брой на лекарите в лечебните и здравните заведение (средно на 10 хил. души)". Жилищната площ в новопостроените жилища в областите нараства средно с 14.6%, при 1‰ по високо ниво на естествения прираст. Съответно жилищната площ в новопостроените жилища в областите нараства средно с 4.7%, при1‰ по високо ниво на механичния прираст. При увеличаване на лекарите в лечебните и здравните заведение (средно на 10 хил. души) с един лекар, жилищната площ при новопостроените жилища в областите нараства средно с 3.2%. Изводите са валидни при елиминирано влияние на промените в средните равнища на жилищната площ в новопостроените жилища поради тренд в динамиката им през изследвания период.

Построените панелни модели не могат да минимизират изцяло пропуските, причинени от невключени фактори, които са специфични за панелните единици (областите). Включването на допълнителни фактори, които влияят върху жилищната площ в новопостроените жилища, също би могло да доведе до по-висока обяснителна способност на иконометричните модели. Може да се очаква, че цената на новопостроените жилища е значим фактор, който определя тяхната динамика. Моделът с фиксирани ефекти би се подобрил значително и с включването на пространствения фактор, поради хетерогенността на цените на новопостроените жилищни имоти в различните области. Потенциалната работа при едни бъдещи изследвания би била именно в тази насока.

Бележки

[1] НСИ: https://www.nsi.bg/sites/default/files/files/metadata/HFund_1_Methodology.pdf

[2] НСИ: https://www.nsi.bg/sites/default/files/files/metadata/Pop_6_Metodology_population.pdf

[3] НСИ: https://www.nsi.bg/sites/default/files/files/metadata/Pop_5_Metodology_migration.pdf

Литература

Атанасова, Цв. (2021) Влияние на демографските промени върху жилищното строителство в област Варна. *Недвижими имоти & бизнес, том 5 (3):* 189-199.

Atanasova, T. (2021) Vliyanie na demografskite promeni varhu zhilishtnoto stroitelstvo v oblast Varna. *Nedvizhimi imoti & biznes, tom 5 (3):* 189-199.

Маринов, Г. (2014) *Панелни единични корени и коинтеграция*. Варна, Издателство: Онгъл.

Marinov, G. (2014) *Panelni edinichni koreni I kointegratsiya*. Varna, Izdatelstvo: Ongal.

Стаматев, Ст. (2018) Отражение на миграционните процеси върху пазара на жилищни имоти в България. *Недвижими имоти & бизнес, том 2 (2)*: 119-127.

Stamatev, St. (2018) Otrazhenie na migratsionnite protsesi varhu pazara na zhilishtni imoti v Balgaria. *Nedvizhimi imoti & biznes, tom 2 (2)*: 119-127.

Томова, Д. (2018) Специфични фактори влияещи върху пазара на жилищни имоти в град София. *Недвижими имоти & бизнес, том 2 (2)*: 103-108.

Tomova, D. (2018) Spetsifichni faktori wliyaeshti varhu pazara na zhilishtni imoti v grad Sofiya. *Nedvizhimi imoti & biznes, tom 2 (2):* 103-108.

Baltagi, B. (2005) *Econometric Analysis of Panel Data, Third Edition*, John Wiley & Sons Ltd.

Greene, W. (2002) *Econometric Analysist*, Fifth Edition.

Gujarti, D., Porter, D. (2009) *Basic Econometrics – 5th edition*. Boston. McGraw-Hill Irwin.

Maddala, G. (2001) *Introduction to Econometrics. Third Edition*, Wiley.

Hsiao, C. (2003) *Analysis of panel data. 2nd edition*. Cambridge University Press.

Wooldridge, J. (2003) *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge.

Assistant Prof. Vesela Dimitrova, PhD
University of National and World Economy, Statistics and Econometrics Department
1700 Sofia, Student Town, UNWE, Office 1089
E-mail: vesela.dimitrova@unwe.bg