

МОДЕЛИ ФИКСНИХ И СЛУЧАЈНИХ ЕФЕКТАТА У АНАЛИЗИ ПОДАТАКА ПАНЕЛА

Апстракт: У овом раду су изложене основне поставке у примјени модела фиксних и случајних ефеката у анализи панел података. Специфичност ових модела је у томе што се прати и временска и структурна компонента података, па се регресиона једначина разликује од опште познате у погледу значења регресионих параметара.

У првом дијелу рада су изложене теоријске поставке, да би се касније прешло на претпоставке од којих се полази у оцјени праметара, како модела са фиксним, тако и оног са случајним ефектима. Иначе, ови модели се примјењују у анализи понашања група домаћинстава, предузећа, земаља итд.

Кључне ријечи: фиксни и случајни ефекти, панел-подаци, оцјене параметара, корелисаност, аутокорелација, временска и структурна компонента, случајна грешка.

1. УВОД

Тема овог рада су фиксни и случајни ефекти у панел-анализи. Специфичност панел-анализе је у томе што она прати и анализира упоредне податке и временске серије истовремено.

У раду су изложене неке основне теоријске поставке у самом приступу овој анализи, поменуте методе које се примјењују, са већим освртом на моделе фиксних и случајних ефеката. Иначе, овакве анализе се, осим у економији, примјењују и у другим истраживањима, као што су социолошка, у истраживањима понашања политичких партија, психологији, здравству итд.

Код овакве анализе, регресиона једначина се разликује у односу на класичну анализу, управо због поменуте двије димензије које посједују подаци (временска и структурна).

Рад је структуриран у неколико дијелова. У првом дијелу се говори уопштено о панел-серијама, њиховим основним карактеристикама, а након тога о општем облику модела панел-података.

Основни дио рада је практично дио који говори о тзв. моделу фиксних учинака. Кажемо основни дио, јер је, у ствари, модел случајних ефеката рјеђе кориштен у пракси и сматра се да је он само одређена корекција модела фиксних ефеката.

Основна специфичност код овог модела је у томе што константни члан у регресионом моделу није исти за све јединице посматрања, он обухвата све варијабле које утичу на зависну варијаблу, а нема их у моделу.

Значајан дио код овог приступа у анализи панел-података посвећен је и оцјени фиксних ефеката, како претпоставкама од којих се полази, неких ограничења која нам се намећу итд.

У последњем дијелу рада говори се о моделу са стохастичким ефектима. Разлика између овог модела и модела фиксних ефеката је у томе што константни члан у регресионом моделу не узима фиксне вриједности, него се понаша као случајна промјенљива. Такође, као и код модела фиксних ефеката и овдје су изложене основне карактеристике модела, као и претпоставке од којих се полази у оцјенама појединих параметара у једначинама преко којих презентујемо овај модел.

2. УОПШТЕНО О ПАНЕЛ-СЕРИЈАМА

Панел-серије су својеврсна комбинација временских серија и упоредних података. Овакве серије садрже податке великог броја јединица (истих) посматрања у одређеном броју временских тачака¹. Ово су, дакле, својеврсне поновљене анкете једног те истог узорка у различитим временским периодима. Велика популарност ових серија лежи у чињеници да свака панел-опсервација носи у себи информације и о структури и о динамици посматране појаве.

Када подаци о некој појави посједују двије димензије (структурну и временску), веома је битно питање како моделирати временску и структурну међузависност изабраних промјенљивих. Да би изабрали одговарајући панел-модел, битно је провјерити претпоставке о хетерогености између јединица посматрања, то јесте о варијабилности регресионих параметара.

Панел-серије углавном садрже податке о појединцима, домаћинствима, предузећима, гранама производње итд. Хетерогеност подразумијева да утицај независних промјенљивих на зависно промјенљиву варира од јединице до јединице посматрања и/или током времена, а и на промјене у самој структури. Варијабилни регресиони параметри указују првенствено на постојање ефеката из модела изостављених индивидуално и/или временских промјенљивих фактора на варијације зависно промјенљиве. Ефекти ових промјенљивих се зову индивидуални и временски ефекти.

2.1. Природа панел-серија

Пошто упоредни подаци представљају упоредне пресјеке, зову се још и серије структуре, односно распореда фреквенција. Серије панел-података односе се на комбинацију упоредних података и временских серија, па оне тако садрже N података истих јединица у T временских тачака.

Битно је напоменути чињеницу да претпоставка о хомогености јединица, која је уобичајена у статистичкој анализи, у економији често није испуњена. Овдје се усваја претпоставка да начин утицаја економских фактора мијења природу и кроз вријеме и по јединицама посматрања.

Регресиона анализа панел серија омогућава не само анализу структуре и хетерогености између јединица посматрања, већ и анализу промјена у структури кроз вријеме.

Један од значајних разлога кориштења панел-серија јесте тај што се тиме значајно повећава узорак и омогућава анализа која не би била изводљива при појединачном кориштењу временских серија или упоредних података. Овим се једноставно обезбјеђује већа ефикасност оцјена параметара модела уз мање рестриктивне претпоставке. Када се понавља проматрање довољног броја група панел-анализа

¹ Радмила С. Драгутиновић Митровић, "Анализа панел серија", Задужбина Андрејевић, 2002.

допушта истраживање динамичности у једном кратком временском низу. Комбинација временског низа са групом управо повећава квалитет и квантитет у смислу да би било немогуће користити само једну димензију посматрања.

За претходно речено могли бисмо искористити следећи примјер као илустрацију: претпоставимо да се оцјењује линеарни регресиони модел, облика: $Y = \alpha + \beta X + u_i$, на основу узорка од N јединица посматрања и T периода.

Када би биле испуњене претпоставке линеарног регресионог модела тада би се методом *OLS* (обичних најмањих квадрата) могло оцијенити T регресионих једначина за N јединица посматрања или N једначина временских серија дужине T .

Тако добијамо оцијењене регресионе параметре за сваки период t ($t=1, \dots, T$) и сваку јединицу посматрања ($i=1, \dots, N$). Овим начином се знатно повећава величина узорка на укупно NT података, чиме се повећава и статистичка ефикасност оцјена регресионих параметара.

Коришћењем ових панел-серија омогућава се анализа ефеката бројних промјенљивих које нису егзактно укључене у модел. Наиме, анализом модела које представљају овакве серије, могу се разликовати индивидуални и временски ефекти које имају промјенљиве, изостављене из модела, на варијације у зависној промјенљивој.

Индивидуални ефекти узимају константне вриједности током времена, али се разликују између јединица посматрања, за разлику од временских ефеката који имају константну вриједност за све јединице посматрања, али се њихова вриједност разликује током времена. Постоје неколике врсте панел-аналитичких метода, а то су: метод константних коефицијената, метод фиксних и случајних ефеката, модели са варијабилним регресионим параметрима уз независне промјенљиве.

Индивидуални и временски ефекти су такође укључени у моделе панел-серија кориштењем вјештачких промјенљивих.

Такође, кориштењем панел-података смањују се економетријски проблеми познати у класичним истраживањима (мултиколинearност између независно промјенљивих, корелација између независно промјенљивих и промјенљивих које нису експлицитно укључене у модел, јер нису уочљиве, или се не могу измјерити па су обухваћене случајном грешком).

Ако су неке од претпоставки модела панел-података нарушене, може доћи до одређених проблема. Тако, на примјер, када су регресиони параметри промјенљиви, али узимају фиксне вриједности у различитим групама посматрања (фиксни утицај), уобичајени метод оцјењивања овог модела компликује оцјењивање услед великог броја параметара. Ако су промјенљиви параметри стохастичког карактера, када случајна грешка нема пожељне особине, тада метод *OLS* даје необјективне оцјене.

Пошто серије панел-података садрже информације о великом броју истих јединица у одређеном великом низу, оне се зову и балансиране или комплетне панел-серије. Дешава се да није могуће обезбиједити податке о истим јединицама посматрања у низу периода, па се такве панел-серије са недостајућим подацима зову небалансиране или некомплетне панел-серије.

Такође некад немамо ситуацију да из периода у период исте јединице посматрања буду у узорку, гдје се за оцјењивање панел-модела умјесто индивидуалних података користе групе јединица посматрања и њихове просјечне вриједности. Ове су серије у литератури познате под именом псеудопанел-серије.

2.2. Општи облик модела панел-података

Предност анализе панел-података је у томе што аналитичару даје велику флексибилност у идентификовању разлика између појединих јединица посматрања². Та се флексибилност управо огледа у могућности кориштења приступа тзв. фиксних и случајних учинака.

Код овог облика регресионог модела панел-података варијације зависне промјенљиве Y објашњене су варијацијама k независно промјенљивих (који је систематски дио) и случајним варијацијама којима је обухваћено дејство промјена других промјенљивих које нису директно укључене у модел (тзв. Стохастички дио модела). Општи облик овог модела изгледа овако:

$$Y_{it} = \beta_{1it}X_{1it} + \beta_{2it}X_{2it} + \dots + \beta_{kit}X_{kit} + u_{it} = \beta_{it} + \sum_{k=2}^k \beta_{kit}X_{kit} + u_{it}$$

$i=1, \dots, N; \quad t=1, \dots, T; \quad k=1, \dots, K,$

гдје је:

Y_{it} - вриједност зависно промјенљиве за i -ту јединицу посматрања у периоду t

X_{kit} - вриједност k -те нез. промјенљиве за i -ту јединицу посматрања у периоду t

$X_{1it} = 1$, за свако i и t

β_{kit} - непознати регресиони параметри који су у општем облику модела панел-података варијабилни по јединицама посматрања у временским периодима.

u_{it} - случајна грешка са аритметичком средином једнаком нули ($E(u_{it})=0$) и конст. заједничком варијансом ($E(u_{it})^2 = \sigma^2_{it}$ за свако i и t).

Из претходно написаног модела видљиво је да постоји различита реакција зависно промјенљиве на варијабле у независним промјенљивим, као и да се она разликује и за сваки временски период. Према томе, свака јединица посматрања има регресиони параметар специфичан за сваки временски период. Оваква општа формулација се не може оцијенити, јер је број регресионих параметара већи од броја података узорка. Да би се овакав модел могао оцјењивати, уводе се претпоставке које се односе на особине регресора (независних промјенљивих) модела, особине случајне грешке, као и на статистичку зависност између регресора и случајне грешке, као и на степен варијабилитета регресионих параметара.

Од скупа претпоставки о хетерогености између јединица посматрања у одређеном периоду, тј. о степену варијабилитета регресионих параметара, зависи избор модела.

2.3. Панел-модел са индивидуалним ефектима

Ако су сви регресиони параметри уз независне промјенљиве константни по јединицама посматрања и у времену ($\beta_{kit} = \beta_k$ за $k=2, \dots, k$), а слободни чланови варијабилни по јединицама посматрања, тада би се оцјењивао модел облика:

$$Y_{it} = \beta_{1i} + \sum_{k=2}^k \beta_k X_{kit} + u_{it}, \quad i=1, \dots, N; \quad t=1, \dots, T; \quad k=1, \dots, K;$$

гдје је:

$\beta_{1i} = \beta_1 + \mu_i$ слободан члан i -те јединице посматрања,

β_1 - просјечна вриједност слободног члана,

² Жељко Ловринчевић, Здравко Марић, Давор Микулић, Приљев инопоземног капитала утјецај на националну штедњу, домаће инвестиције и биланцу плаћања транзиционих земаља Средње и Источне Европе'

u_i – величина одступања β_{1i} од просјека β_1 . (ова величина у ствари представља ефекте из модела изостављених промјенљивих које узимају константну вриједност током времена, али се разликују по јединицама посматрања).

Овдје се претпоставља да слободни члан не варира по различитим периодима за дату јединицу посматрања, тј. $\beta_{1i} = \beta_{1i}$. Уколико је испуњена претпоставка да су слободни чланови β_{1i} промјенљиви са фиксним вриједностима, код различитих јединица посматрања, тј. када је u_i фиксни параметар, говори се о моделу фиксних ефеката. Када се индивидуални ефекти посматрају као стохастичке промјенљиве, онда се ради о моделу са стохастичким ефектима.

3. МОДЕЛ ФИКСНИХ ИНДИВИДУАЛНИХ ЕФЕКТА

3.1. Уопштено о моделу фиксних ефеката

Овај модел се заснива на претпоставци да су индивидуални ефекти који су обухваћени слободним члановима, непознати и фиксни параметри. Он се још зове и панел-модел са вјештачким варијаблама. Код фиксних учинака у обзир се узима специфичност сваке поједине групе посматрања. Тако константни члан у регресионом моделу није заједнички за све јединице посматрања, а обједињује све варијабле које могу утицати на зависну варијаблу, а које се изостављене из модела. Константни члан се тако тумачи као вриједност зависне промјенљиве, ако је вриједност независне промјенљиве нула. Суштина методе фиксних ефеката може се сажети и рећи да једноставно свака јединица посматрања служи сама себи као контрола. То се постиже тако што се праве поређења унутар јединица посматрања и након тога се просјечна вриједност тих разлика примјени на све јединице у узорку.

Метод фиксних ефеката у потпуности занемарује варијације између јединица посматрања и усредсређује се само на варијације унутар њих. Одбацивање варијација између јединица посматрања може довести до стандардних грешака које су знатно веће од оних које настају користећи методе са варијацијама и унутар и између објеката посматрања.

Користећи метод фиксних ефеката могуће је имати контролу над свим карактеристикама посматраних јединица. Уколико је зависна варијабла квантитативно изражена, онда се овај метод може лако примијенити користећи обичну линеарну регресију најмањих квадрата. Два су кључна предуслова за примјену методе фиксних ефеката:

1. свака јединица посматрања у узорку мора имати више мјерења у односу на исту зависну варијаблу.
2. за бар неке од јединица у узорку вриједности независних варијабли морају бити различите при најмање два мјерења.

3.2. Оцјена фиксних ефеката

Код оцјена модела (како фиксних, тако и случајних ефеката), користећи OLS, може доћи до значајних компликација због малог броја опсервација по јединицама посматрања (оцјене нису конзистентне). Потешкоће се односе на то што је временско кашњење зависне варијабле повезано са термином грешке, чак и кад сами поремећаји нису аутокорелисани. Да би се отклонили ови проблеми, Арелјано и Бонд су развили

оцјену генерализованог метода момената, који рјешава претходно поменуте тешкоће. Оцјена динамичких панел-података Арелјано и Бонда користи следеће инструменте: нивое зависне варијабле која "касни" два или више периода, нивое ендогених варијабли које "касне" два или више периода и првог извода строго егзогених коваријата који се користе као сопствени инструменти. Да бисмо пратили метод оцјене ових ефеката, можемо поћи од следеће релације: $Y_{it} = \beta_{0i} + \beta_1 X_{it} + a_i + u_{it}$

За сваку поједину опсервацију i , једначина за просјечне вриједности у сваком тренутку је:

$$\bar{Y}_i = \beta_{0i} + \beta_1 \bar{X}_i + a_i + \bar{u}_i$$

Одузимајући просјечну вриједност зависно промјенљиве од вриједности зависно промјенљиве у било којем временском тренутку, имали бисмо:

$$Y_{it} - \bar{Y}_i = \beta_1 (X_{it} - \bar{X}_i) + u_{it} + \bar{u}_i, \quad t=1, \dots, T;$$

Односно, краће написано:

$$\dot{Y}_{it} = \beta_1 \dot{X}_{it} + \dot{u}_{it}, \quad t=1, \dots, T;$$

Гдје је $\dot{Y}_{it} = Y_{it} - \bar{Y}_i$ и

$$\dot{X}_{it} = X_{it} - \bar{X}_i,$$

$$\dot{u}_{it} = u_{it} - \bar{u}_i.$$

За разлику од овог модела, у моделу прве диференце објашњавајуће варијабле, које су константне цијело вријеме, не могу се укључити јер је $\dot{X}_{it}=0$.

Иако се временски константне варијабле не могу укључивати саме у модел фиксних ефеката, оне се могу укључивати у интеракцији са варијаблама које се мијењају кроз вријеме - са тзв. бинарним варијаблама. Исто тако, у модел је немогуће укључити и објашњавајуће варијабле чије су промјене константне у времену.

Уколико имамо више објашњавајућих варијабли, једначина ће изгледати овако:

$$Y_{it} = \beta_0 + \beta_1 X_{it1} + \beta_2 X_{it2} + \dots + \beta_k X_{itk} + a_i + u_{it} \text{ или}$$

$$\dot{Y}_{it} = \beta_1 \dot{X}_{it1} + \beta_2 \dot{X}_{it2} + \dots + \beta_k \dot{X}_{itk} + \dot{u}_{it}$$

Претпоставке непристрасних оцјена:

1. Дат је нпр. модел: $Y_{it} = \beta_1 X_{it1} + \dots + \beta_k X_{itk} + a_i + u_{it}$
2. Случајни узорак је узорак упоредних података или нпр. међусекторске димензије итд.
3. За свако $t \in (u_{it} | X_{it}; a_i) = 0$
4. Свака објашњавајућа варијабла се мијења кроз вријеме и између објашњавајућих варијабли нема корелације
5. За свако $t \quad \text{var}(u_{it} | X_{it}; a_i) = \text{var}(u_{it}) = \sigma^2_{it}$
6. За $t \neq s, \text{cov}(u_{it}, u_{is} | X_{it}; a_i) = 0$

Оцјене у моделу фиксних ефеката су ефикасније од оцјена у моделу прве диференце.

Бинарне варијабле могу се интерпретирати са a_i (ознака из претходних једначина), наравно за сваку опсервацију i .

Ако бисмо рачунали за сваку појединачно i , онда би се, користећи модел фиксних ефеката, $\hat{a}_i$ добијало на следећи начин:

$$\hat{a}_i = \bar{Y}_i - \hat{\beta}_1 \bar{X}_{i1} - \hat{\beta}_2 \bar{X}_{i2} - \dots - \hat{\beta}_k \bar{X}_{ik}, \quad i=1, \dots, N;$$

3.2.1. Прва диференца наспрам фиксних ефеката

*За $T=2$:

Оцјене су и у једном и у другом моделу, као и статистички тестови, идентичне.

*За $T \geq 3$:

У оба модела оцјене су непристрасне и конзистентне, тако кад је N велико и T мало, основна рјешења су релативно ефикаснија. (претпоставка хомоскедастичности)

Ако грешке у времену слиједе случајни ход, тако да је Δu_{it} у корелацији, онда је I диференца више ефикаснија.

Када је T велико и N мало:

- оцјене у моделу фиксних ефеката су веома осјетљивије на нарушавање претпоставки класичног модела фиксних ефеката, прва диференца се показује бољом.
- Међутим, фиксни ефекти, мање су осјетљиви на нарушавање стриктне егзогености.

Наставља се коришћење уобичајеног броја временских периода који се користе у опсервацијама упоредних података.

Неке опсервације са само једним периодом не играју улогу у анализи фиксних ефеката јер бисмо имали: $\dot{Y} = 0, \dot{X} = 0$

Ако је у овом типу модела присутна ситуација да су подаци у корелацији са случајном грешком, оцјене су пристрасне.

Међутим, ако су подаци у корелацији са фиксним ефектима α_i , онда то неће резултирати пристрасношћу.

Као и остали модели, ни модели фиксних ефеката нису без недостатака. Примјена овог модела може захтијевати превише варијабли, што може ослабити модел довољног броја степени слободе за примјену адекватно моћног статистичког теста. Осим тога, овакав модел може бити оптерећен мултиколинеарношћу, са повећањем стандардних грешки у оцјењивању параметара. Такође, иако је претпоставка да су грешке модела нормално дистрибуиране и хомогене, лако може доћи до појаве хетероскедастичности или аутокорелације.

4. МОДЕЛ СА СТОХАСТИЧКИМ ИНДИВИДУАЛНИМ ЕФЕКТИМА

4.1. Уопштено о моделу стохастичких ефеката

Ако су варијабилни слободни чланови β_{it} стохастичке промјенљиве, а не фиксни параметри, тада важе следеће претпоставке:

$$\beta_{kit} = \beta_k, \text{ за свако } i, t \text{ и } k=2, \dots, K$$

$$\beta_{1i} = \beta_1 + \mu_i, \text{ односно } \mu_i + \mu_{it} = v_{it}$$

$$u_{it} \sim N(0, \sigma_u^2); u_i \sim N(0, \sigma_\mu^2)$$

Вратимо се накратко на модел:

$$Y_{it} = \beta_{1i} + \sum_{k=2}^K \beta_k X_{kit} + u_{it}, \quad i=1, \dots, N; \quad t=1, \dots, T; \quad k=1, \dots, K;$$

Ако се слободни члан β_{1i} третира као случајна промјенљива, онда се горњи модел зове модел са стохастичким индивидуалним ефектима, односно модел са компонентама случајне грешке. Претпоставка од које се овде полази је да је компонента слободног члана u_i стохастичка промјенљива ($\beta_{1i} = \beta_1 + \mu_i$), док је компонента β_1 непознати

параметар који представља средњу вриједност слободних чланова у цијелом основном скупу.

дати сада поприма облик:

$$Y_{it} = \beta_1 + \mu_i + \sum_{k=2}^k \beta_k X_{kit} + u_{it} = \beta_1 + \sum_{k=2}^k \beta_k X_{kit} + v_{it}$$

гдје је $v_{it} = \mu_i + u_{it}$

Сам термин "компоненте случајне грешке" потиче из структуре случајне грешке v_{it} . Ова случајна грешка садржи двије компоненте, и то:

- случајну промјенљиву μ_i .
- остатак грешке u_{it} .

4.2. Оцјена случајних ефеката

И избор метода оцјењивања модела са компонентом случајне грешке зависи од својстава случајне грешке v_{it} . Претпоставке о случајној грешци v_{it} су:

1. $E(v_{it})=0$, значи, очекивана вриједност случајне грешке једнака је нули.
2. $var(v_{it}) = var(\mu_i) + var(u_{it}) = \sigma_\mu^2 + \sigma_u^2$.
3. $cov(v_{it}, v_{js}) = \sigma_\mu^2$ за $i=j$, $t \neq s$
4. $cov(v_{it}, v_{js}) = 0$ за $i \neq j$, $t=s$

Видимо, дакле, да су случајне грешке у различитим периодима за i -ту јединицу посматрања корелисане, док случајне грешке различитих јединица посматрања нису корелисане.

Једноставно, у моделу стохастичких ефеката, индивидулни ефекти су обухваћени случајном грешком v_{it} . То значи да хетерогеност међу јединицама посматрања није директно укључена у регресиони модел помоћу очекиване вриједности зависне промјенљиве, као што је то било код модела фиксних ефеката, већ је хетерогеност обухваћена њеном варијансом.

Литература:

1. Радмила С. Драгутиновић Митровић
"Анализа панел серија", Задужбина Андрејевић, 2002.
2. Green, W.H., "Econometric Analysis"
5th ed., Prentice Hall, 2003.
3. Милена Јовичић, "Економетријски методи"
Економски факултет Београд, 2002.
4. Damodar N. Gujarati, "Basic Econometrics"
4th ed., McGraw-Hill, 2003.
5. <http://www.nyu.edu>

Младен Фулерија

FIXED AND RANSOM EFFECTS MODELS IN THE PANEL DATA ANALYSIS

SUMMARY

This paper exposes the basic postulates in application of fixed and ransom effect models in the analysis of the panel data. The special characteristic of these models is in that you can follow the time and structural components of the data, so the regression equation differs from the common one regarding the meaning of the regressive parameters.

The first part of the paper exposes the theoretic postulates, continuing with starting assumptions for parameters assessment, both of models with fixed and ransom effects. These models are applied in the analysis of the groups of households, enterprises, countries etc.

Key words: fixed and ransom effects, panel data, parameters assessment, correlate, autocorrelation, time and structural components, accidental mistake.