

Analýza panelových údajov – aplikácia na odvetvia hospodárstva Slovenskej republiky

Michal BENČÍK*

Úvod

Hospodársky výsledok (zisk) predstavuje časť vytvoreného (hrubého domáceho) produktu, patriacu podnikateľským subjektom. V našich podmienkach s pomerne nízkymi úsporami domácností by to mal byť dôležitý finančný zdroj na reštrukturalizáciu podnikovej sféry. Výšku hospodárskeho výsledku ovplyvňujú mnohé faktory platné len pre individuálny podnik. Zo zložiek hospodárskeho výsledku má asi najväčšiu pravidelnosť hospodársky výsledok z finančných operácií, pretože existuje trh úverov, na ktorom sú určité podmienky spoločné pre všetky subjekty. Navyše, v podmienkach Slovenska treba v rámci analýzy hospodárskeho výsledku venovať zvýšenú pozornosť hospodárskemu výsledku z finančných operácií, keďže strata z finančných operácií je často taká veľká (alebo väčšia) ako súčet ostatných (kladných) zložiek hospodárskeho výsledku. Strata z finančných operácií teda často zapríčiňuje aj celkovú stratu podniku alebo odvetvia.

Cieľom tohto článku je ozrejmiť metodológiu práce s panelovými údajmi a aplikovať ju pri odvetvovej analýze hospodárskeho výsledku z finančných operácií.

Príspevok má dve relatívne samostatné časti. V prvej časti popisujeme štruktúru použitých údajov a metódu ich spracovania a v druhej časti popísanú metódu po predbežnej prezentácii údajov aplikujeme.

1. Metodológia práce s panelovými údajmi

1.1 Charakteristika panelových údajov

Na štatistickú analýzu je potrebné jednotlivé údajové položky (pozorovania) systematicky zoradiť. Pri ekonometrickom výskume môžeme používať tri hlavné typy údajov: *časové rady*, *prierezové údaje* a *panelové údaje*.

* Ing. Michal BENČÍK, CSc., Ústav slovenskej a svetovej ekonomiky SAV, Šancová 56, 811 05 Bratislava 1

Za zapožičanie literatúry a za konzultácie ďakujem Mgr. Gabriele Moravčíkovej z Ústavu aplikovanej matematiky MFUK v Bratislave.

Časové rady

Časové rady sú hodnoty istého ukazovateľa za tú istú entitu merané v rovnakých časových intervaloch, napríklad ročné hodnoty hrubého domáceho produktu (HDP) za Slovenskú republiku. Tieto údaje sa pri ekonometrickej analýze používajú najčastejšie. Vo väčšine prípadov medzi skúmanými veličinami existujú pomerne stabilné vzťahy, ktoré pri regresnej analýze vedú k dobrému vyrovnaniu. K tomu prispieva aj skutočnosť, že rušivé vplyvy okolia sa často dajú z tvaru rezíduí určiť a ošetriť umelými premennými. Regresné rovnice z časových radov sa často používajú na prognózovanie.

Pri časových radoch z makroekonomických veličín je najčastejším problémom autokorelácia rezíduí. Tým je porušený predpoklad Gaussovej-Markovovej vety, že variančno-kovariančná matica náhodných zložiek je skalárny násobok jednotkovej matice, pretože tá má nenulové prvky aj mimo diagonály. V prípade autokorelácie nie sú hodnoty odhadov jednoduchou metódou najmenších štvorcov (JMNS) výdatné a ak medzi vysvetľujúcimi premennými (regresormi) je časovo posunutá závisle premenná (regresand), nie sú ani konzistentné. Tento problém sa najčastejšie rieši zmenou dynamickej špecifikácie príslušnej rovnice.

V poslednom čase sa začínajú konštruovať modely, v ktorých sa predpokladá, že rozptyl náhodných zložiek nie je konštantný, ale závisí od rozptylu predchádzajúcej náhodnej zložky, prípadne aj iných premenných (modely autoregresívne podmienené heteroskedastických – ARCH – procesov; definíciu pozri napr. v [2]). Teraz má teda variančno-kovariančná matica náhodných zložiek na diagonále rôzne prvky. Táto konštrukcia vyplýva z poznatku, že na kapitálových trhoch sa striedajú obdobia relatívneho pokoja s obdobiami veľkých výkyvov. Tieto modely sa konštruujú pre ceny akcií, burzové indexy a výmenné kurzy z údajov s krátkou periodicitou (nanajvýš mesačnou). V našich podmienkach však zrejme nie sú ešte tieto modely použiteľné.

Prierezové údaje

Prierezové údaje predstavujú súbor pozorovaní pre rôzne entity toho istého druhu v rovnakom čase, napríklad index rastu HDP v krajinách sveta za obdobie 1970 až 1990 (podiel hodnôt za tieto dva roky). Tieto údaje sa modelujú oveľa ťažšie, pretože jednotlivé štáty majú viac špecifik oproti teoretickým vzťahom ako jednotlivé obdobia. Často sú údaje a informácie, ktoré by viedli k vysvetleniu nepredvídaných odchýlok, nedostupné alebo nekvantifikovateľné. Preto sa niekedy pri regresiach s prierezovými údajmi považuje za uspokojivú taká hodnota koeficientu determinácie, ktorá implikuje zamietnutie nulovej hypotézy o nezávislosti regresandu od množiny regresorov (F -test). Preto má regresná analýza pri prierezových údajoch za cieľ obvykle len analýzu vzťahov medzi veličinami ex post.

Pri prierezových údajoch je najdôležitejším problémom heteroskedasticita, ktorá tiež ruší výdatnosť odhadov jednoduchou metódou najmenších štvorcov. V danej súvislosti táto vlastnosť vyplýva z toho, že jednotlivé analyzované entity nie sú celkom porovnateľné, ale odlišujú sa napríklad veľkosťou. Tento problém sa rieši buď použitím intenzitných ukazovateľov (produktivita na pracovníka, kapitálová vybavenosť pracovného miesta), alebo sa skonstruuje špeciálny model pre rozptyly náhodných zložiek a predikcia tejto matice sa zohľadní pri odhade parametrov. Tieto postupy sa môžu aj kombinovať.

Panelové údaje

Panelové údaje sú pozorovania za viac porovnateľných entít merané vo viacerých rovnako dlhých intervaloch, napríklad ročné hodnoty HDP členských štátov EÚ za roky 1970 až 1990 (za každý rok jedna hodnota pre každý štát). Panelové údaje majú teda vlastnosti tak časových radov, ako aj prierezových údajov. Na rozdiel od týchto typov však majú dva indexy – jeden pre jednotlivé subjekty a druhý pre čas.¹ Podľa toho, či je v súbore málo entít pre veľa období, alebo veľa entít s niekoľkými pozorovaniami, prevažujú vlastnosti buď časových radov, alebo prierezových údajov. Na základe toho treba testovať a odstraňovať aj prípadné porušenia predpokladov Gaussovej-Markovovej vety.

Panelové údaje vyžadujú osobitné metódy hlavne kvôli špecifickému spôsobu, ako sú v nich zohľadnené informácie. Táto situácia je do istej miery analogická predpokladom analýzy rozptylu. Pri analýze rozptylu celkový súbor rozdelíme na viac skupín – podsúborov. Z tých potom vypočítame priemer a rozptyl. Rozptyl celkového súboru sa dá rozložiť na súčet rozptylov jednotlivých skupín a na rozptyl priemerov týchto skupín okolo priemeru celkového súboru. Ak je posledný člen oproti súčtu rozptylov jednotlivých skupín malý, jednotlivé skupiny sú si podobné a celkový súbor je homogénny. Ak je rozptyl skupinových priemerov okolo celkového priemeru pomerne veľký, znamená to, že jednotlivé skupiny – podsúbory sa od seba značne odlišujú. Na takomto princípe sa zakladá aj F -test, ktorým sa testuje homogenita súborov.

Panelové údaje môžeme rozdeliť napríklad tak, že vznikne toľko vektorov prierezových údajov (skupín), koľko je časových období. Predpokladu homogeneity je potom analogický predpoklad stability parametrov v čase. Samotné rozdiely medzi skupinami môžu byť tiež funkciami stochastických veličín. Tieto skutočnosti sú zohľadnené v špeciálnych metódach, ktoré ďalej popíšeme.

¹ Od tohto typu údajov treba odlišiť prípad, keď sa pri prierezových údajoch používajú premenné z rôznych období. Napríklad Menbere Workie Tiruneh [6] pri štúdiu konvergencie krajín používa na vysvetlenie logaritmického priemernej miery rastu krajín okrem iných premenných aj logaritmus počítačovej úrovne outputu, ale celá analýza je jednorozmerná.

1.2 Typy odhadov z panelových údajov, ich vlastnosti a použitie

Uvažujme nasledujúci prípad: Máme skúmať regresnou analýzou údaje, kde všetky premenné sú organizované do panelov. Každý panel má údaje za N rôznych entít (index i od 1 po N) a T časových období (index t od 1 po T), a teda obsahuje $T \times N$ pozorovaní. Ďalej, so zreteľom na naše údaje, uvažujme, že N je značne väčšie ako T . Znamená to, že jednotlivé panely (matice) môžeme rozdeliť na T vektorov prierezových údajov, každý s N zložkami. V nasledujúcom výklade budú tieto vektory zodpovedať skupinám. Ďalej predpokladajme, že ich budeme vysvetľovať K premennými, z ktorých jedna bude lokujúca konštanta.

Najjednoduchší odhad získame tak, že z pozorovaní urobíme vektor s $T \times N$ pozorovaniami a odhadneme parametre rovnice jednoduchou metódou najmenších štvorcov, akoby mali len jeden rozmer:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix} = \begin{bmatrix} e \\ e \\ \vdots \\ e \end{bmatrix} \alpha + \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_T \end{bmatrix} \beta + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{bmatrix}$$

kde y_1 až y_T sú vektory závisle premennej s rozmerom $N \times 1$ pre obdobia – skupiny 1 až T , e je súčtový vektor rozmeru $N \times 1$, α^* je odhad lokujúcej konštanty, X_1 až X_T sú matice vysvetľujúcich premenných bez konštanty rozmeru $N \times K-1$, β je vektor regresných koeficientov rozmeru $K-1 \times 1$ a u_1 až u_T sú vektory náhodných zložiek rozmeru $N \times 1$.

Takto získame odhad jednoduchou metódou najmenších štvorcov (*OLS estimator*), ktorý predpokladá konštantný vzťah medzi skúmanými veličinami pri rôznych entitách a v rôznych časových obdobiach. Ak urobíme analógiu našej regresnej analýzy s analýzou rozptylu, toto je prípad, keď parametre v jednotlivých skupinách sú rovnaké a variabilita medzi skupinami je spôsobená variabilitou jednotlivých pozorovaní (preto aj názov *between group estimator* [3]). Z uvedeného vyplýva, že tento odhad je skreslený, keď sa jednotlivé parametre menia v závislosti od skupín alebo času (*heterogeneity bias*). Ignoruje totiž druhý rozmer údajov.

Najjednoduchšou možnosťou, ako pri odhade zohľadniť povahu panelových údajov, je odhad lokujúcej konštanty pre každé obdobie (skupinu)² osobitne. Pri tom vznikne z konštanty sada T umelých premenných. Tie majú všade nulovú

² Teraz je dôležité, ako sme si zvolili jednotlivé skupiny. V našich údajoch máme okolo štyridsať odvetví a len štyri roky. Z praktického hľadiska je oveľa lepšie ak robíme štyri regresie (pre každý rok jednu), každú zo štyridsiatich pozorovaní, než keby sme robili štyridsať regresii (pre každé odvetvie jednu), každú zo štyroch pozorovaní.

hodnotu, len v pozorovaniach zodpovedajúcich t -tej skupine majú jednotky. Tento odhad s umelými premennými sa volá aj *least squares dummy variable* (LSDV) estimator. Na rozdiel od predošlého tento odhad považuje skupiny za pomerne samostatné celky, variabilitu medzi skupinami vysvetľuje rozdielnosťou parametrov *v jednotlivých skupinách* a na odhad využíva len variabilitu v rámci skupín (odtiaľ názov *within-group estimator* [3]). Tento typ odhadu teda predpokladá, že parametre merajúce vplyv vysvetľujúcich premenných na závisle premennú sú v čase rozdielne, ale nestochastické (čiže *fixed effects*). Uvažujme s údajmi v už popísanom tvare. Tento odhad je výsledkom modelu

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix} = \begin{bmatrix} e \\ 0 \\ \vdots \\ 0 \end{bmatrix} \alpha_1^* + \begin{bmatrix} 0 \\ e \\ \vdots \\ 0 \end{bmatrix} \alpha_2^* \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ e \end{bmatrix} \alpha_T^* + \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_T \end{bmatrix} \beta + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{bmatrix}$$

kde 0 je nulový vektor rozmeru $N \times 1$, α_1^* až α_T^* sú odhady lokujúcej konštanty pre jednotlivé obdobia a ostatné symboly znamenajú to isté ako v predchádzajúcom vzorci.

Možno získať aj samostatné odhady podľa jednotlivých skupín. Tieto odhady sa dajú spojiť do jedného výpočtu, no vysvetľujúce premenné sa nepospájajú do tvaru vektorov, ale do blokovo diagonálnej matice. Všetky parametre môžu byť rôzne pre jednotlivé obdobia – skupiny (pozri [1]):

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix} = \begin{bmatrix} e & X_1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & e & X_2 & & \vdots \\ \vdots & & & & \ddots & 0 & 0 \\ 0 & \cdots & & 0 & e & X_T \end{bmatrix} \beta + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{bmatrix}$$

kde vektor β má TK zložiek (sú v ňom aj odhady lokujúcich konštant) a vysvetľujúce premenné sú usporiadané do blokovo diagonálnej matice s rozmermi $TN \times TK$. Tento odhad je však oproti ostatným pomerne málo výdatný.

Odhad JMNS a odhad s umelými premennými predstavujú dva extrémny. Kým prvý počíta len s variabilitou údajov medzi skupinami, druhý predpokladá, že variabilita je nulová. Kompromisom medzi nimi je vážený priemer týchto dvoch odhadov. Ten sa obvykle interpretuje pomocou stochastických efektov. Parametre teda majú pravdepodobnostné rozdelenie a ich variabilitu meria príslušný rozptyl. Od tohto rozptylu a reziduálneho rozptylu závisí váha odhadov uvedených v odhade parametrov pri predpoklade stochastických (náhodných) efektov. Výsledný odhad by mal byť neskreslený (na rozdiel od odhadu JMNS) a výdatnejší ako odhad s umelými premennými. Postup výpočtu prezentuje Ch. Hsiao [3].

Nech I je jednotková matica rozmeru N a celá matica na ľavej strane nasledujúcej rovnice má rozmer $N(T-1) \times NT$, vektor β má TK zložiek a nulový vektor 0 má $N(T-1)$ zložiek. Keby sme na všetkých TK zložkách vektora β „naložili“ ohraňenie v tvare

$$\begin{bmatrix} I & -I & & \dots & 0 \\ & I & -I & & \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & & I & -I \end{bmatrix} \beta = 0$$

z odhadu s diagonálnou maticou dostaneme odhad JMNŠ (vektor β bude mať T rovnakých sekcií [1]). Tento tvar možno v prípade heteroskedasticity náhodnej zložky normalizovať nasledujúcim spôsobom:

1. odhadneme variančno-kovariančnú maticu náhodnej zložky $\hat{\Sigma}$;
2. vypočítame maticu $\hat{\Sigma}^{-1/2}$, kde $\hat{\Sigma}^{-1/2} \hat{\Sigma}^{-1/2} \hat{\Sigma} = I_{NK}$ (jednotková matica rozmeru NT);
3. maticou $\hat{\Sigma}^{-1/2}$ zľava vynásobíme závisle premennú i maticu vysvetľujúcich premenných.

Takto sa získa homoskedastická náhodná zložka, ktorá je nutnou podmienkou testovania vedľajších ohraňení.

1.3 Postup použitý v tomto príspevku

V našom prípade máme do činenia s heteroskedasticitou. Opatrenia na jej odstránenie ovplyvňujú odhady reziduálneho rozptylu, čo spôsobuje určité metodologické problémy pri výpočte váh pre odhad s náhodnými efektmi. Preto sme použili odhad s blokovo-diagonálnou maticou. Korigovali sme však heteroskedasticitu a ohraňili sme parametre s tým istým znamienkom a prekryvajúcimi sa konfidenčnými intervalmi.

Heteroskedasticitu sme testovali Whiteovým testom [7]. Štvorce rezíduí z odhadu s blokovo-diagonálnou maticou sme vysvetlili tými istými regresormi, ich druhými mocninami a niektorými súčinmi. Údaje sme usporiadali do blokovo-diagonálnej matice podobne ako pri odhade regresných koeficientov. Testovacou štatistikou bol súčin počtu pozorovaní a koeficientu determinácie. Potom sme vypočítali p -hodnotu príslušného χ^2 rozdelenia, a keďže bola menšia ako 5 %, prijali sme hypotézu, že reziduá sú heteroskedastické. Vypočítali sme podobnú regresiu ako pri teste, ale ako nezávisle premennú sme použili logaritmus štvorca odchýlok. Vyrovnané hodnoty z tejto regresie, ktoré zodpovedali teoretickým rozptylom náhodných zložiek, sme odlogaritmovali, odmocnili a prevrátili.

Takto sme získali váhy (maticu $\hat{\Sigma}^{-1/2}$). Všetky premenné z odhadu s blokovo-diagonálnou maticou sme vynásobili týmito váhami. Transformovaný model mal potom homoskedastickú náhodnú zložku.

Parametre tohto modelu sme potom ohraničili vedľajšími podmienkami. Ale ak sú medzi jednotlivými obdobiami – skupinami rozdiely, potom sú odhady s ohraničením skreslené. Preto sme pri našom odhade použili len tie riadky z uvedenej matice ohraničení, ktoré zodpovedali koeficientom s približne rovnakými hodnotami. Ak určíme, že dva parametre sa rovnajú a ony sa v skutočnosti od seba nelíšia, je to tak, akoby sme odhadovali len jeden parameter. Ak sme použili pri odhade šesťnásť takých ohraničení, „ušetrili“ sme šesťnásť parametrov bez výrazného zvýšenia reziduálneho rozptylu. Takto sme dosiahli väčšiu výdatnosť nášho odhadu.

2. Analýza hospodárskeho výsledku z finančných operácií

2.1 Použité údaje

Dáta na túto analýzu sme čerpali z podkladov dodaných bankovým zúčtovacím centrom (BZC). Obsahujú záznamy o viacerých položkách zo štatistických výkazov *Súvaha* a *Výsledovka* za celé odvetvia v dvojmiestnej klasifikácii OKEČ (za každé odvetvie osobitne), ako aj počty podnikov za roky 1994 až 1997 (s výnimkou odvetví, kde pre malý počet podnikov nebolo podľa Zákona o ochrane dát možné poskytnúť ani odvetvové agregáty). Obdobnú databázu za roky 1993–1996 sme mali k dispozícii aj minulý rok. Bankové zúčtovacie centrum však priebežne doplnilo údaje za všetky obdobia, a tak sa údaje za roky 1994 až 1996 v starej a novej verzii databázy od seba odlišujú. Preto sme sa rozhodli použiť len údaje z novej verzie.

Databázu z BZC sme porovnávali s údajmi, ktoré publikuje Štatistický úrad Slovenskej republiky a Národná banka Slovenska. Položka *Výnosy* z databázy BZC tvorí asi 85 % ukazovateľa *Hrubá produkcia* z národných účtov. Zhruba porovnateľné sú súčty položiek *Bankové úvery a výpomoci* z databázy BZC s korunovými úvermi podnikom a domácnostiam, ako ich publikuje Národná banka Slovenska.

2.2 Grafická analýza

Hospodársky výsledok podnikov sa skladá z troch zložiek: hospodársky výsledok z prevádzkovej činnosti, hospodársky výsledok z finančných operácií a mimoriadny hospodársky výsledok. Hospodársky výsledok z prevádzkovej činnosti

je určený faktormi špecifickými pre dané odvetvie, hlavne relatívnymi cenami, technológiou použitou v danom odvetví a konkurenciou zo zahraničia. Mimo-riadny hospodársky výsledok zahŕňa inštitucionálne a mimoriadne vplyvy. Tieto zložky sú špecifické pre každé odvetvie, a preto nie je vhodné tieto ukazovatele porovnávať medzi odvetviami. Hospodársky výsledok z prevádzkovej činnosti sa však riadi pravidlami, ktoré sa dajú zovšeobecniť na všetky nefinančné odvetvia. Finančné operácie sa totiž riadia pravidlami, ktoré sú pre všetky nefinančné organizácie do veľkej miery rovnaké. Preto má zmysel hľadať v tvorbe hospodárskeho výsledku z finančných operácií pravidlá spoločné pre všetky odvetvia. Podiel hospodárskeho výsledku z finančných operácií bude teda naša závisle premenná.

Vzhľadom na sektorové špecifiká sa môže postavenie jednotlivých odvetví na trhu s úvermi odlišovať. Preto sme súbor odvetví rozdelili do sektorov:

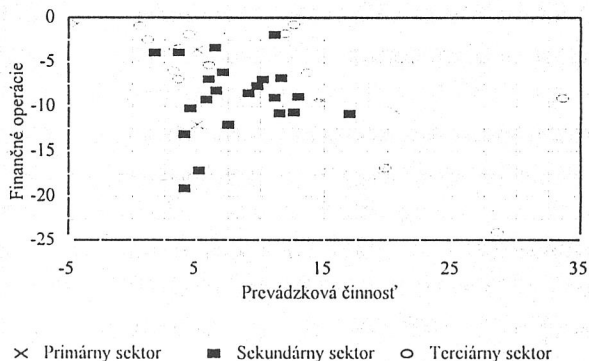
- primárny sektor – odvetvia OKEČ 01 až 10;
- sekundárny sektor – odvetvia OKEČ 15 až 37 (odvetvia 11 až 14 chýbajú);
- terciárny sektor – odvetvia OKEČ 40 až 73 (odvetvia 38 a 39 chýbajú).

Údaje za netrhové služby, štátnu správu, zdravotníctvo, školstvo atď. sme síce mali k dispozícii, ale predpokladáme natoľko odlišné fungovanie týchto odvetví, že by to našu analýzu skreslilo. Preto sme ich neanalyzovali. Keďže chýbali údaje aj za iné odvetvia, analyzovali sme asi spolu 40 odvetví za každý rok.

Na rozptylových grafoch 1–4 sú znázornené prvé dve zložky hospodárskeho výsledku, vydelené výnosmi za jednotlivé roky. Vo všeobecnosti je zrejma tendencia odvetví s vyšším prevádzkovým ziskom dosahovať vyššiu stratu z finančných operácií. V každom roku sa však dá identifikovať skupina priemyselných odvetví, ktoré majú nepriaznivé hodnoty v oboch ukazovateľoch – tento stav je pravdepodobne výsledkom dlhodobej kumulácie problémov.

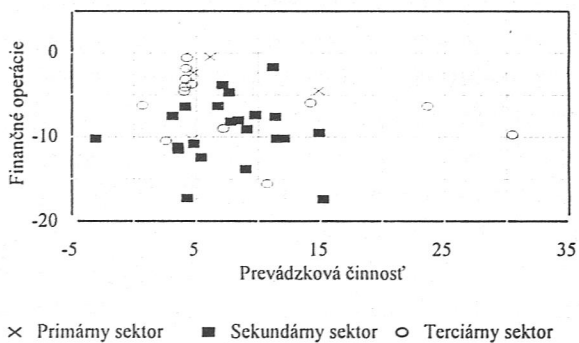
Graf 1

Vzťah medzi zložkami rentability 1994 – bez bankovníctva



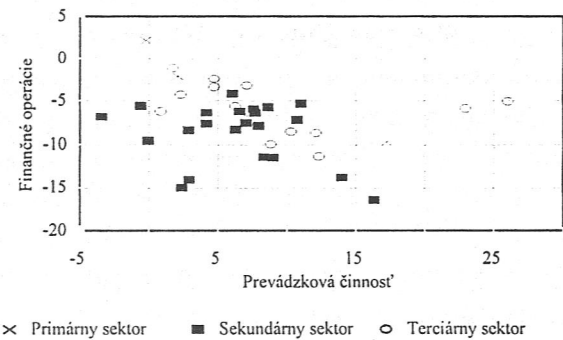
Graf 2

Vzťah medzi zložkami rentability 1995 – bez OKEČ 65, 67 a 70



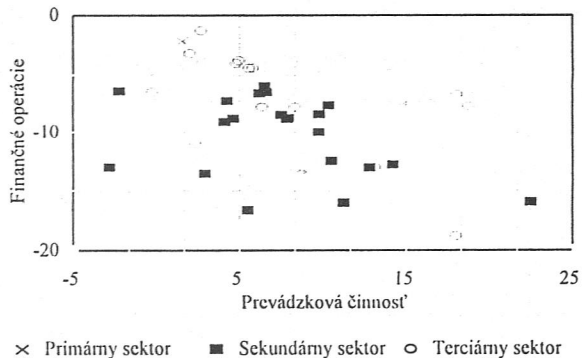
Graf 3

Vzťah medzi zložkami rentability 1996 – bez OKEČ 65 a 67



Graf 4

Vzťah medzi zložkami rentability 1997 – bez OKEČ 65 a 67



2.3 Regresná analýza

Hospodársky výsledok z finančných operácií budeme vysvetľovať konštantami (pre každý rok jedna), podielom úverov na výnosoch, podielom zadlženosti na výnosoch a podielom hospodárskeho výsledku na výnosoch. Údaje sme navyše dezagregovali podľa sektorov – uvedené grafy indikujú rozdiely medzi sektormi. Finančné organizácie a iné odľahlé pozorovania³ sme podobne ako v minulej verzii vyradili, aj keď ich nebolo toľko ako predtým. Extenzitné ukazovatele vedú k modelu s heteroskedastickou náhodnou zložkou, okrem toho sú ovplyvnené veľkosťou odvetvia, a to môže tiež spôsobiť koreláciu veličín, ktoré spolu nesúvisia. Preto sme dáta zhomogenizovali tak, že používame podiely príslušných veličín na výnosoch. Rozptyl náhodnej zložky však môže naďalej závisieť napríklad od regresorov, preto aj v modeli po úprave testujeme heteroskedasticitu, a keďže je prítomná, berieme na ňu zreteľ.

Použité údaje majú osobitnú štruktúru, pretože majú dva rozmery: odvetvie a rok. Natíska sa otázka, či vzťahy medzi hospodárskym výsledkom z finančných operácií a ostatnými veličinami sú stále v čase a medzi sektormi.⁴ Stabilita v tomto prípade znamená rovnosť parametrov pri premenných z jednotlivých rokov a sektorov. Po aplikácii metód popísaných v predchádzajúcej časti sme dostali tento výsledok (použili sme ekonometrický program SORITEC 6.5 pre DOS).

Závislá premenná WGREVF1 je vážený a pospájaný rad podielu hospodárskeho výsledku z finančných operácií na výnosoch. Nezávislých premenných je veľa, preto sú označené nasledujúcim spôsobom:

- predpona WG označuje vážené a nulami doplnené rady;
- nasledujúca skupina znakov vyjadruje:
 - CONST – konštanta,
 - DLHV – podiel zadlženosti na výnosoch,
 - FMV – podiel finančného majetku na výnosoch,
 - UVV – podiel bankových úverov a výpomocí na výnosoch,
 - REVP – podiel hospodárskeho výsledku z prevádzkovej činnosti na výnosoch,
 - DUMMY3 – umelá premenná pre tretí sektor popísaná vyššie;
- dvojčíslice 94 až 97 znamená rok;
- číslica 1, 2 alebo 3 za rokom (ak tam je) predstavuje primárny, sekundárny a terciárny sektor (pre ostatné sektory má nulovú hodnotu).

³ Odľahlé pozorovania sa svojimi hodnotami značne líšia od zvyšku súboru [5].

⁴ Doteraz sme uvažovali so štyrmi skupinami – pre každý rok jednu. Teraz sme však pri niektorých premenných odvetvia v jednotlivých rokoch rozdelili do sektorov, ako sme to popísali v časti 2.2, takže sme dostali dvanásť skupín (štyri roky krát tri sektory). Potom sme testovali rovnosť nielen medzi zodpovedajúcimi parametrami z rôznych období, ale aj z rôznych sektorov.

RLS: dependent variables is WGREVF1

USE period has many gaps.

Restricted Least-Squares

Variable	Coefficient	Std Err	T-stat	Signf
WGCONST94	-1.04528	0.360866	-2.89660	0.004
WGDLHV941	-6.27086	1.36596	-4.59080	0.000
WGDLHV942	-2.35940	1.16409	-2.02682	0.045
WGDLHV943	-2.35940	1.16409	-2.02682	0.045
WGFMV94	4.46206	2.15243	2.07303	0.040
WGUVV941	-32.4471	1.81913	-17.8366	0.000
WGUVV942	-32.4471	1.81913	-17.8366	0.000
WGUVV943	-32.4471	1.81913	-17.8366	0.000
WGREVP942	-0.841466E-01	0.177606E-01	-4.73784	0.000
WGCONST95	-4.02294	0.679379	-5.92151	0.000
WGDLHV951	-8.95232	1.51839	-5.89593	0.000
WGDLHV952	1.38548	1.25842	1.10096	0.273
WGDLHV953	1.38548	1.25842	1.10096	0.273
WGFMV95	45.8784	6.78759	6.75917	0.000
WGFMV953	-6.67983	4.91839	-1.35813	0.177
WGUVV951	-14.9450	5.02381	-2.97484	0.004
WGUVV952	-31.9920	2.66853	-11.9886	0.000
WGUVV953	-31.9920	2.66853	-11.9886	0.000
WGREVP95	-0.841466E-01	0.177606E-01	-4.73784	0.000
WGCONST96	-1.04528	0.360866	-2.89660	0.004
WGDLHV961	-8.95232	1.51839	-5.89593	0.000
WGDLHV962	-2.35940	1.16409	-2.02682	0.045
WGDLHV963	1.38548	1.25842	1.10096	0.273
WGFMV96	13.8470	5.09025	2.72030	0.008
WGFMV963	-32.5698	6.00400	-5.42469	0.000
WGUVV961	-13.9271	1.72664	-8.06605	0.000
WGUVV962	-26.3053	1.48028	-17.7705	0.000
WGUVV963	-13.9271	1.72664	-8.06605	0.000
WGREVP96	-0.185618	0.350621E-01	-5.29398	0.000
WGCONST97	-2.61572	0.441452	-5.92527	0.000
WGDLHV971	1.38548	1.25842	1.10096	0.273
WGDLHV972	-13.8003	3.41934	-4.03596	0.000
WGDLHV973	-2.35940	1.16409	-2.02682	0.045
WGFMV97	23.0518	6.92282	3.32983	0.001
WGFMV973	-6.67983	4.91839	-1.35813	0.177
WGUVV971	-6.91948	2.91681	-2.37228	0.019
WGUVV972	-21.1472	2.93767	-7.19864	0.000
WGUVV973	-21.1472	2.93767	-7.19864	0.000
WGREVP97	-0.185618	0.350621E-01	-5.29398	0.000

R-Squared = 0.992677 No. obs = 157

R-Bar-Squared (Adj for df) = 0.991420

Log of likelihood function = -281.969

Durbin-Watson (45 gaps) = 1.934715

Durbin-Watson (4) (45 gaps) = 2.081943

Sum of squared residuals = 333.729

Std. error of regression = 1.57814

Sum of residuals = -0.658408

Mean of dependent variable = -9.54631

F-statistic (23, 134) = 789.724

Significance level = 0.000000

Hypothesis SSR = 16.7156

df = 16 Mean Sq = 1.04473

F (16, 118) = 0.388873

Significance = 0.982928

Parametre pri jednotlivých premenných možno teda interpretovať ako jednotkové finančné náklady, resp. výnosy jednotlivých faktorov. Parametre pri hospodárskom výsledku z prevádzkovej činnosti však treba pri porovnávaní vynásobiť stomi, pretože táto premenná bola v percentách a ostatné nie. Použitá metóda nám umožnila testovať rovnosť parametrov v čase a v jednotlivých sektoroch. Parametre, ktoré sa od seba štatisticky neodlišujú, majú rovnaké hodnoty aj štandardné odchýlky. Tieto hodnoty však treba hodnotiť skôr ako priemerné.

V sledovanom období stratu z finančných operácií najpravidelnejšie ovplyvňovali bankové úvery a výpomoci. Najviac sa za ne platilo v roku 1994 – asi 32 hal. na korunu úverov. Platby za úvery sa v ďalších rokoch znižovali, v niektorých sektoroch aj veľmi výrazne. V roku 1995 sa v sekundárnom a terciárnom sektore platilo 31 hal. na korunu úverov, čo sa od predchádzajúceho roka nelíši, ale v primárnom sektore to bolo oveľa menej – asi 15 hal. Tento pokles však nemožno posudzovať pozitívne. Ak jednotkové platby klesnú pod úroveň priemernej úrokovej sadzby z úverov, znamená to, že príslušný sektor si v príslušnom roku v priemere neplatí ani úroky. To znamená, že buď samovoľne narastá masa (zlých) úverov, alebo banky nejakú časť úverov odpisujú. V roku 1996 mal okrem primárneho sektora nízke jednotkové platby (asi 14 hal. na korunu úverov) aj terciárny sektor. V roku 1997 klesli jednotkové platby primárneho sektora až na 7 hal. Treba však podotknúť, že veľkosť týchto parametrov sa mení podľa toho, aké ďalšie premenné sú do modelu zahrnuté a aké ohraničenia musia splňať príslušné parametre.

Hospodársky výsledok z prevádzkovej činnosti je, ako už preukázala grafická analýza, späť aj so stratou z finančných operácií. Táto veličina na rozdiel od ostatných zachytáva skôr výrobnotechnické než finančné rozhodnutia. Preto sme ju do predchádzajúcej verzie (1998) ani nezaradili. So stratou z finančných operácií však zjavne súvisí. Jej vplyv v čase rastie – v roku 1994 jednej korune zisku z prevádzkovej činnosti zodpovedalo 8 hal. straty z finančných operácií v sekundárnom sektore, v roku 1995 sa jej vplyv potvrdil už vo všetkých sektoroch a v posledných dvoch rokoch sa tento pomer zvýšil asi na dvojnásobok. Hoci podľa rozptylových grafov i korelačných matíc mal v sekundárnom sektore zisk z prevádzkovej činnosti vplývať na stratu z finančných operácií pozitívne, po rozšírení analýzy o ďalšie premenné sa sekundárny sektor v tomto ohľade prestal odlišovať od ostatných sektorov. Sú dve možnosti, ako zápornú koreláciu v tomto prípade vysvetliť: jednak ako všeobecné platby za peniaze, jednak ako výrobný faktor (hlavne v tých odvetviach, v ktorých je podiel vlastných zdrojov na celkových zdrojoch malý). Túto hypotézu potvrdzuje aj kladná korelácia medzi bankovými úvermi a ziskom z prevádzkovej činnosti, hlavne v sekundárnom sektore. Pokiaľ by to však bol jediný dôvod, zisk z prevádzkovej činnosti

by mal byť nevýznamný a jeho vplyv by mali „prebrať“ bankové úvery a výpomoci. Iná možnosť je, že čím viac zisku z prevádzkovej činnosti podniky vyprodukujú, tým viac sa ho snaží manažment podnikov preliať cez finančné operácie do iných subjektov.

Zadlženosť (rozdiel medzi záväzkami a pohľadávkami) vplyva na stratu z finančných operácií pomerne slabo. Nielenže sú jednotkové náklady na tieto zdroje nižšie, ale v niektorých prípadoch zadlženosť stratu z finančných operácií znižuje – kladné znamienko sa vyskytuje v sekundárnom a terciárnom sektore v roku 1995, v terciárnom sektore v roku 1996 a v primárnom sektore v roku 1997. Znamená to, že týmto sektorom sa darilo prenášať svoje problémy na svojich dodávateľov neplatením.

Najnepravidelnejší vplyv na zisk z finančných operácií mal finančný majetok. V roku 1994 zarobili podniky 4,5 hal. na korune finančného majetku za rok (čiže 4,5 %). V roku 1995 primárny a sekundárny sektor dosiahli zisk vyše 45 hal. na korunu finančného majetku za rok, terciárny sektor mal výnos iba 39 hal. za rok (toto číslo sme dostali sčítaním príslušných parametrov 45,6 a -6,8). Už v nasledujúcom roku sa však pomer výnosov obrátil – primárny a sekundárny sektor mali jednotkový výnos asi 13 %, terciárny sektor mal stratu asi 18 %. V roku 1997 sa v podnikovej sfére situácia do istej miery zlepšila: zatiaľ čo primárny a sekundárny sektor dosiahli výrazný zisk (asi 23 %), terciárny sektor zaznamenal nižší zisk (asi 18 %). Vo všetkých rokoch okrem roka 1994 sa teda potvrdila odlišnosť výnosnosti portfólia terciárneho sektora od portfólia primárneho a sekundárneho sektora. Terciárny sektor mal v rokoch 1995–1997 významne nižšie výnosy z finančného majetku, držal teda iné cenné papiere ako ostatné dva sektory. Aby sme však mohli tento jav vysvetliť, potrebovali by sme dodatočné informácie, ktoré teraz k dispozícii nemáme. Parametre však jasne ukazujú na heterogenitu finančného majetku v podnikoch.

Záver

Tento príspevok sa zaoberá aplikáciou metodológie analýzy panelových údajov na hospodársky výsledok z finančných operácií. Panelové údaje sú charakteristické tým, že každá premenná tvorí samostatnú pseudomaticu (vektor s viacerými analogickými sekciami). Osobitosť týchto metód spočíva v delení tohto vektora pozorovaní na sekcie a v skúmaní hodnôt parametrov pre jednotlivé sekcie. Ak sa predpokladá rovnosť parametrov pre všetky sekcie, odhad je totožný s jednoduchou metódou najmenších štvorcov (JMŇS) aplikovanou na všetky pozorovania. Druhým extrémom je model fixných efektov, v ktorom sa odhady JMŇS aplikujú na každú sekciiu osobitne. Kompromisom medzi týmito

prístupmi je model náhodných efektov, ktorý považuje parametre za funkcie náhodnej premennej a z matematického hľadiska je to vážený priemer oboch predošlých odhadov. My sme však efektívnosť odhadu fixných efektov zvýšili uložením vedľajších podmienok na parametre modelu, čím sme zohľadnili rozdielnu premenlivosť jednotlivých parametrov v čase.

Podľa vykonaných výpočtov stratu z finančných operácií spôsobovali najmä pasívne úroky z úverov. Náklady asociované so zadlženosťou sa nepotvrdili vo všetkých sektoroch. Všeobecne sa však potvrdila väzba straty z finančných operácií so ziskom z prevádzkovej činnosti. V primárnom a sekundárnom sektore výnosy z finančného majetku stratu znižovali, no v terciárnom sektore sa v roku 1996 dosahuje strata spojená s finančným majetkom a výnosy v ostatných rokoch sú tiež nižšie. Keďže ani vývoj výnosov finančného majetku primárneho a sekundárneho sektora nekopíruje vývoj výnosov terciárneho sektora, finančný majetok primárneho a sekundárneho sektora asi tvoria iné aktíva ako finančný majetok terciárneho sektora.

Došlo 21. 2. 2000

Literatúra

- [1] BALTAGI, B. H.: *Econometric Analysis of Panel Data*. London, John Willey and Sons 1995.
- [2] HENDRY, D. F.: *Dynamic Econometrics*. Oxford etc., Oxford University Press 1995.
- [3] HSIAO, Ch.: *Analysis of Panel Data*. Cambridge, Cambridge University Press 1986.
- [4] JURÍČKOVÁ, V. a kol.: *Mikroekonomická rovnováha a rozpočtová politika štátu*. [Expertízna štúdia, č. 24.] Bratislava, Ústav slovenskej a svetovej ekonomiky SAV 1999.
- [5] MELOUN, M. – MILITKÝ, J.: *Statistické zpracování experimentálních dat na osobních počítačích*. Pardubice, Finish 1992.
- [6] MENBERE, W. T.: *Economic Growth and Convergence in a Cross-Section in Economics*. *Ekonomický časopis/Journal of Economics*, 46, 1998, č. 5, s. 632–654.
- [7] RAMANATHAN, R.: *Introductory Econometrics with Applications*. Third edition. Fort Worth etc., The Dryden Press 1995.

PANEL DATA ANALYSIS APPLIED TO BRANCHES OF SLOVAK ECONOMY

Michal BENČÍK

This article presents the methodology of panel data analysis and applies it analyzing profits/losses resulting from financial operations in various branches in Slovak economy.

Panel data is characterized by two dimensions for every variable. One of them is time (corresponding index t is going from 1 to T) and the other one are entities (index j is going from 1 to N). Like in conventional regression analysis, we wish to explain the dependent

variable (regressand) with K independent variables (regressors; index i is going from 1 to K), first of them being the constant term. The data matrices can be split into smaller groups, we are going to split the data in T groups, where every group contains observations from different entities and fixed time (in a group, index t is fixed).

The simplest estimator in this setup is the OLS estimator. It is a $K \times 1$ vector β . It is obtained by simply pooling all groups together into $NK \times 1$ vectors and applying the conventional ordinary least squares method. It corresponds to the hypothesis, that there are no significant differences between groups – if this does not hold, OLS estimator will be biased.

One can segregate the first independent variable, the summation vector corresponding to the constant term, into T dummies, with every dummy corresponding to one group. By doing so, the constant term is allowed to vary across the groups – the resulting LSDV estimator consists of T estimates of the constant term (for every group a different estimate) and $(K - 1) \times 1$ vector β corresponding to other regressors. This estimator is likely to be less efficient, because it uses only a fraction of the information contained in the data.

The disadvantages of above estimators led to construction of random effects estimator. It assumes that the parameters themselves are random variables and leads to an estimator in similar form as LSDV estimator, but the estimates of constant term are nearer to the constant from OLS estimator. Some efficiency should be gained.

In general, not only constant terms vary across the groups; one can let all parameter vary. In this case, the independent variables are arranged into a block-diagonal matrix. The parameter vector has dimension $TK \times 1$ and consists of T sections at K parameters each. This is the most general approach, but it is the least efficient one. To increase its efficiency one can use restrictions imposing equality of parameters across different groups. In the limiting case, if all parameters are constant between groups, we arrive at the simple OLS estimator (its values are repeated in every group). To test these restrictions one usually has to transform the model to a homoskedastic one, as heteroskedasticity is usually present and use a F -test for linear restrictions. We prefer gradually adding and testing equality constraints to construction of random-effects estimator, because the hypothesis is clearer and it is a data-driven approach.

In our application, we used data from Bank Clearing Center in the form of branch aggregates of some balance sheet and income sheet for years 1994–1997. The graphic analysis revealed a negative correlation between profits/loss from financial operations and EBIT. As the behaviour of enterprises can vary across primary, secondary and tertiary sector, we spliced the big groups further into sectors for some variables.

We used the panel data methodology to explain profits/loss from financial operations with excess debt (difference between liabilities and claims), financial assets, bank loans and EBIT. All variables are divided with total revenues in order to adjust to branch size. We proceed with construction of the estimator with block-diagonal matrix of regressors and made it more parsimonious by restrictions across years and sectors.

The most regular impact on the dependent variable exercise bank loans. Their impact culminated in 1994 (0.35 SKK on 1 SKK principal) and decreased in time. In some sectors, the average payments fell even under average interest rate, indicating cumulation of bad debts in banks.

The impact of EBIT was smaller in 1994 (parameter value 0.08), but it increased over time. According to scatter graph, in secondary sector it should be higher, but after other explanatory variables were added, it was not. In general, however, the results support the hypothesis, that enterprises tend to „smuggle“ profits via financial operations in other subjects.

Excess debt (not paying one's suppliers) is in our conditions much cheaper than loans (0.02 to 0.12 SKK for 1 SKK debt) and in some sectors, it is even positive, indicating that these enterprises shift their problems to their suppliers.

According to our analysis, return on financial assets are in all periods except 1994 higher in primary and secondary sectors than in the tertiary sector. The return in primary and secondary sectors approaches lower values in 1994 and 1996, slightly higher values in 1995 and 1997. The difference between them and tertiary sector culminates in 1996, when the tertiary sector suffers even loss. Thus, we can conclude, that the main determinant of profits or loss from financial operations were bank loans and in primary and secondary sectors in years 1995 and 1997 also financial assets. The relation between of profits or loss from financial operations and EBIT was established, but it was quite weak, the same can be said also about excess debt.