

UNIVERZITA PARDUBICE
FAKULTA EKONOMICKO-SPRÁVNÍ

DIPLOMOVÁ PRÁCE

2024

Bc. Marcela Fillová

Univerzita Pardubice
Fakulta ekonomicko-správní

Využití analýzy panelových dat v podnikové praxi

Diplomová práce

Univerzita Pardubice
Fakulta ekonomicko-správní
Akademický rok: 2023/2024

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Bc. Marcela Fillová**
Osobní číslo: **E22519**
Studijní program: **N0413A050009 Ekonomika a management**
Specializace: **Ekonomika a management podniku**
Téma práce: **Využití analýzy panelových dat v podnikové praxi**
Zadávající katedra: **Ústav matematiky a kvantitativních metod**

Zásady pro vypracování

V práci budou popsány vybrané modely používané při analýze panelových dat. Cílem práce pak bude aplikovat analýzu panelových dat na konkrétní problém z podnikové praxe, zvolit vhodný model, odhadnout a interpretovat jeho parametry.

Osnova:

- Základy analýzy panelových dat.
- Konstrukce vybraných modelů panelových dat.
- Výběr optimálního modelu.
- Analýza výsledků.

Rozsah pracovní zprávy: **50**
Rozsah grafických prací:
Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam doporučené literatury:

ARLT, Josef a Markéta ARLTOVÁ. *Ekonomické časové řady*. Praha: Professional Publishing, 2009. ISBN 978-80-86946-85-6.
BALTAGI, Badi H. *Econometric Analysis of Panel Data*. 6th ed. Cham: Springer Cham, 2021. ISBN 978-3-030-53952-8.
CIPRA, Tomáš. *Finanční ekonometrie*. 2., upr. vyd. Praha: Ekopress, 2013. ISBN 978-80-86929-93-4.
WOOLDRIDGE, Jeffrey M. *Introductory Econometrics: A Modern Approach*. 3rd ed. Mason: Thomson South-Western, 2006. ISBN 0-324-32348-4.

Vedoucí diplomové práce: **Mgr. Hana Boháčová, Ph.D.**
Ústav matematiky a kvantitativních metod

Datum zadání diplomové práce: **1. září 2023**
Termín odevzdání diplomové práce: **30. dubna 2024**

prof. Ing. Jan Stejskal, Ph.D. v.r.
děkan

L.S.

doc. Ing. et Ing. Renáta Myšková, Ph.D. v.r.
garant studijního programu

V Pardubicích dne 1. září 2023

Prohlašuji:

Práci s názvem Využití analýzy panelových dat v podnikové praxi jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Beru na vědomí, že v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů, a směrnicí Univerzity Pardubice č. 7/2019 Pravidla pro odevzdávání, zveřejňování a formální úpravu závěrečných prací, ve znění pozdějších dodatků, bude práce zveřejněna prostřednictvím Digitální knihovny Univerzity Pardubice.

V Pardubicích dne 26. 6. 2024

Bc. Marcela Fillová v. r.

PODĚKOVÁNÍ

Na tomto místě bych ráda poděkovala vedoucí této diplomové práce paní Mgr. Haně Boháčové, Ph.D. za cenné rady a čas, který mi během tvorby této práce věnovala.

ANOTACE

Tato diplomová práce se zaměřuje na problematiku týkající se analýzy panelových dat. Popisuje jednotlivé modely panelových dat. Aplikuje analýzu panelových dat na konkrétní problém z podnikové praxe. Součástí práce je výběr vhodného modelu, jeho aplikace a interpretace odhadnutých parametrů.

KLÍČOVÁ SLOVA

panelová data, analýza, modely s individuálními efekty, dynamický model

TITLE

Use of panel data analysis in business practice.

ANNOTATION

This thesis focuses on issues related to panel data analysis. It describes different models of panel data. It applies panel data analysis to a specific problem in business practice. The thesis includes the selection of an appropriate model, its application and interpretation of the estimated parameters.

KEYWORDS

panel data, analysis, individual effects models, dynamic model

OBSAH

ÚVOD.....	12
1. Základy analýzy panelových dat.....	13
1.1 Vývoj panelových dat	13
1.2 Vymezení panelových dat.....	14
1.2.1 Panel.....	18
1.3 Výhody a problémy panelových dat	19
1.4 Testy jednotkových kořenů.....	21
1.5 Oblasti využití panelových dat	22
1.6 Panelový výzkum.....	22
1.6.1 Příklady panelových výzkumů.....	24
2. Konstrukce vybraných modelů panelových dat.....	27
2.1 Panelová regrese	27
2.2 Spojený regresní model	28
2.3 Modely s individuálními efekty	29
2.3.1 Model s fixními efekty	30
2.3.2 Model s náhodnými efekty.....	31
2.4 Dynamické modely	32
2.4.1 Autoregresivní modely.....	33
3. Aplikace analýzy panelových dat	34
3.1 Charakteristika dat	34
3.2 Popisné statistiky	34
3.3 Aplikace vybraných modelů na jednorozměrná panelová data	37
3.3.1 Spojený regresní model.....	37
3.3.2 Model s fixními efekty	39
3.3.3 Model s náhodnými efekty.....	42
3.4 Výběr optimálního modelu	45
3.4.1 Analýza výsledků vybraného modelu	45
3.5 Aplikace vybraných modelů na vícerozměrná panelová data	51
3.5.1 Spojený regresní model.....	51
3.5.2 Model s fixními efekty	52
3.5.3 Model s náhodnými efekty.....	56

3.6	Výběr optimálního modelu	59
3.6.1	Analýza výsledků vybraného modelu	59
3.7	Závěrečné porovnání a vyhodnocení	64
3.8	Doporučení	66
	ZÁVĚR	67
	POUŽITÁ LITERATURA	69
	SEZNAM PŘÍLOH	72

SEZNAM OBRÁZKŮ

Obrázek 1 Spojený regresní model 1	38
Obrázek 2 Model s fixními efekty 1	39
Obrázek 3 Model s fixními efekty se zahrnutím individuálních časových efektů 1.....	41
Obrázek 4 Model s náhodnými efekty 1	43
Obrázek 5 Model s náhodnými efekty se zahrnutím individuálních časových efektů 1	44
Obrázek 6 Spojený regresní model 2.....	52
Obrázek 7 Model s fixními efekty 2	53
Obrázek 8 Model s fixními efekty se zahrnutím individuálních časových efektů 2.....	55
Obrázek 9 Model s náhodnými efekty 2.....	56
Obrázek 10 Model s náhodnými efekty se zahrnutím individuálních časových efektů 2	58

SEZNAM TABULEK

Tabulka 1 Typy statistických analýz	14
Tabulka 2 Zobrazení struktury panelových dat	15
Tabulka 3 Dlouhý formát panelových dat	16
Tabulka 4 Široký formát panelových dat.....	17
Tabulka 5 Popisné statistiky proměnných	35
Tabulka 6 Popisné statistiky jednotlivých stánků – tržby.....	35
Tabulka 7 Popisné statistiky jednotlivých stánků – návštěvnost.....	35
Tabulka 8 Popisné statistiky jednotlivých stánků – počet zaměstnanců	36
Tabulka 9 Průměrné hodnoty proměnných.....	37
Tabulka 10 Součty čtverců reziduí pro jednotlivé modely	45
Tabulka 11 Koeficienty individuálních časových efektů	46

Tabulka 12 Přehled celkových tržeb jednotlivých zápasů.....	47
Tabulka 13 Hodnoty individuálních časových efektů v rámci zápasů se stejným soupeřem 1	48
Tabulka 14 Hodnoty individuálních časových efektů v rámci zápasů se stejným časem začátku utkání 1	49
Tabulka 15 Hodnoty individuálních časových efektů v rámci zápasů se stejným dnem konání 1	50
Tabulka 16 Konstanty vztažené k jednotkám 1	50
Tabulka 17 Porovnání průměrných tržeb a konstant 1	51
Tabulka 18 Součty čtverců reziduí 2	59
Tabulka 19 Koeficienty individuálních časových efektů 2	60
Tabulka 20 Hodnoty individuálních časových efektů v rámci zápasů se stejným soupeřem 2	61
Tabulka 21 Hodnoty individuálních časových efektů v rámci zápasů se stejným časem začátku utkání 2	62
Tabulka 22 Hodnoty individuálních časových efektů v rámci zápasů se stejným dnem konání 2	63
Tabulka 23 Konstanty vztažené k jednotkám 2	63
Tabulka 24 Porovnání průměrných tržeb a konstant 2	64
Tabulka 25 Rozdíly mezi průměrnou tržbou a konstantami.....	65

ÚVOD

Analýza panelových dat nachází využití v nejrůznějších oblastech například v sociologii, makroekonomii, lékařství, demografii nebo v podnikové praxi. Na poslední zmiňovanou oblast se zaměřuje tato diplomová práce.

V dnešní podnikatelské sféře je pro firmy velmi důležité disponovat nástroji, které jim pomohou lépe porozumět chování zákazníků, efektivněji cílit své marketingové kampaně, lépe nastavit své firemní strategie nebo zefektivnit interní procesy a dosáhnout tak lepších obchodních výsledků. Analýza panelových dat je v tomto ohledu cenným nástrojem, jelikož umožňuje sledovat chování jednotlivců či entit v čase a odhalovat skutečnosti, které by z průřezových dat nebo časových řad nebyly zřejmé.

Panelová data, někdy též označovaná jako longitudinální, představují soubor dat, který sleduje charakteristiky stejných jednotek v průběhu více časových období. Tato data umožňují sledovat vývoj sledovaných veličin a zkoumat kauzální vztahy mezi nimi. Jedná se o komplexní datové soubory, které obsahují velké množství pozorování. Jejich získávání je pro firmy časově i finančně náročné. Vzhledem k náročnosti jejich zpracování zaznamenala analýza panelových dat výraznější rozvoj až v posledních desetiletích, a to díky technologickému rozvoji a lepší dostupnosti softwarů pro analýzu dat.

Cílem této diplomové práce je aplikovat analýzu panelových dat na konkrétní problém z oblasti podnikové praxe. Práce se zaměří na analýzu závislosti tržeb na návštěvnosti a počtu zaměstnanců stánků s občerstvením na fotbalovém stadionu. Na základě poskytnutých panelových dat bude vybrán vhodný model a budou odhadnuty jeho parametry. Odhadnuté parametry budou poté interpretovány, následně bude navrženo doporučení pro provozovatele stánků. Součástí práce bude teoretické vymezení základů panelových dat a představení jednotlivých modelů panelových dat, které mohou být při jejich analýze využity.

Tato diplomová práce by měla poskytnout základní teoretický přehled o panelových datech, především se ale zaměřuje na praktickou aplikaci modelů panelových dat za použití dat z podnikové praxe.

1. ZÁKLADY ANALÝZY PANELOVÝCH DAT

Analýza panelových dat se řadí mezi kvantitativní statisticko-ekonometrické metody zpracování dat. Pomáhá pochopit změny, vzorce a trendy v chování proměnných. Sleduje, jak ovlivňují změny jedné proměnné chování ostatních proměnných, tím určuje kauzalitu vztahů. Analýza panelových dat je také schopna predikovat budoucí vývoj jednotlivých proměnných.

1.1 Vývoj panelových dat

Nerlove¹ ve své práci uvádí, že počátek vývoje analýzy panelových dat začal již v roce 1861, kdy George Biddell Airy zpracovával astronomická data pomocí modelu s náhodnými efekty. K rozvoji analýzy panelových dat také přispěl na začátku 20. století Eugen Fisher, když studoval lidskou dědičnost.

Výraznější rozvoj ale analýza panelových dat zaznamenala až v druhé polovině 20. století. Vzhledem k náročnosti zpracování se totiž rozvoj analýzy panelových dat pojí s rozvojem technologií. V této době se v USA také začali realizovat jedny z prvních panelových výzkumů – National Longitudinal Survey (1966) a Panel Study of Income Dynamics (1968).² První aplikace panelových dat v rámci ekonometrie proběhla v roce 1962, kdy Irving Hoch odhadoval parametry Cobb-Douglasovy produkční křivky. Analyzoval tak data amerických farem v průběhu několika let.³

V roce 1977 se konala v Paříži první mezinárodní konference, jejímž cílem je sdílení poznatků expertů v oblasti panelových dat. Od té doby proběhlo již 28 ročníků, další konference se bude pořádat v červenci 2024.⁴

¹ NERLOVE, Mark. *Essays in Panel Data Econometrics*. Cambridge University Press, 2002. s. 6. ISBN 978-0-521-815345-5.

² HSIAO, Cheng. *Analysis of Panel Data*. 4th ed. Cambridge University Press, 2022. s. 1. ISBN 978-1-316-51210-4

³ HOCH, Irving. Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data. Online. 1962. Dostupné z: <https://www.jstor.org/stable/1911286>. [cit. 2024-06-17].

⁴ *About us*. Online. 29th International Panel Data Conference. 2024. Dostupné z: <https://ipdc2024.org>. [cit. 2024-06-17].

1.2 Vymezení panelových dat

Cipra⁵ rozlišuje v oblasti ekonometrie tři základní skupiny dat, a to časová, průřezová a panelová data.

Časová data jsou vyjádřena formou časových řad. Jedná se o hodnoty, které jsou pozorovány v určitém intervalu s určitou frekvencí záznamu a jsou zaznamenávány chronologicky. Tyto hodnoty nemusí být zaznamenávány pravidelně. Naproti tomu **průřezová data** jsou hodnoty pozorované v jeden časový okamžik skrz určitý datový soubor. Tyto hodnoty nemusí být zaznamenávány v určitém pořadí.⁶

Panelová data jsou souborem časových řad pro jednotlivé statistické jednotky, kde jsou sledované statistické znaky měřeny opakovaně v čase.⁷ Zvolené statistické jednotky se nazývají panelem a v čase se nesmějí měnit, jinak by se jednalo pouze o sdružená průřezová data. Lze tedy říct, že panelová data, někdy nazývaná jako longitudinální, jsou kombinací průřezových dat a časových řad.⁸

Panelová data tedy obsahují tři dimenze – dimenzi pozorovaných objektů, dimenzi sledovaných proměnných a časovou dimenzi. Hebák⁹ uvádí přehled jednotlivých typů statistických analýz (viz Tabulka 1).

Tabulka 1 Typy statistických analýz

typ statistické analýzy	počet objektů	počet proměnných	počet časových okamžiků
jednorozměrná statistická analýza	více	jedna	jeden
jednorozměrná analýza časové řady	jeden	jedna	více
jednorozměrná analýza panelových dat	více	jedna	více
vícerozměrná statistická analýza	více	více	jeden
vícerozměrná časová řada	jeden	více	více
vícerozměrná analýza panelových dat	více	více	více

Zdroj: vlastní zpracování dle Hebák, 2005

⁵ CIPRA, Tomáš. *Finanční ekonometrie*. 2., upr. vyd. Praha: Ekopress, 2013. s. ISBN 978-80-86929-93-4.

⁶ ARLT, Josef a ARLTOVÁ, Markéta. *Ekonomické časové řady*. Praha: Professional Publishing, 2009. s. 12. ISBN 978-80-86946-85-6.

⁷ HSIAO, Cheng. *Analysis of Panel Data*. 4th ed. Cambridge University Press, 2022. ISBN 978-1-316-51210-4.

⁸ CIPRA, Tomáš. *Finanční ekonometrie*. 2., upr. vyd. Praha: Ekopress, 2013. ISBN 978-80-86929-93-4.

⁹ HEBÁK, Petr; HUSTOPECKÝ, Jiří; JAROŠOVÁ, Eva a MALÁ, Ivana. *Vícerozměrné statistické metody*. Praha: Informatorium, 2005. s. 27. ISBN 80-7333-025-3.

Struktura panelových dat lze značit různými způsoby. Tato práce bude používat následující značení:

y_{it} je vysvětlovaná proměnná jednotky i ($i = 1, 2, \dots, n$) a čase t ($t = 1, 2, \dots, T$),

x_{itj} je hodnota j -té vysvětlující proměnné ($j = 1, 2, \dots, k$) pro i -tou jednotku v časovém okamžiku t ,

u_{it} je reziduální složka pro jednotku i v časovém okamžiku t .

V případě, že by platilo $T = 1$, jednalo by se o průřezová data. Pokud by se počet objektů rovnal jedné, šlo by o časové řady.

Tabulka 2 ukazuje možnou strukturu panelových dat.

Tabulka 2 Zobrazení struktury panelových dat

objekt	čas	závisle proměnná	nezávisle proměnná			
i	t	y_{it}	x_{itj}			
1	1	y_{11}	x_{111}	x_{112}	...	x_{11k}
1	2	y_{12}	x_{121}	x_{122}	...	x_{12k}
...
1	$T-1$	y_{1T-1}	x_{1T-11}	x_{1T-12}	...	x_{1T-1k}
1	T	y_{1T}	x_{1T1}	x_{1T2}	...	x_{1Tk}
2	1	y_{21}	x_{211}	x_{212}	...	x_{21k}
2	2	y_{22}	x_{221}	x_{222}	...	x_{22k}
...
2	$T-1$	y_{2T-1}	x_{2T-11}	x_{2T-12}	...	x_{2T-1k}
2	T	y_{2T}	x_{2T1}	x_{2T2}	...	x_{2Tk}
...
$n-1$	$T-1$	y_{n-1T-1}	$x_{n-1T-11}$	$x_{n-1T-12}$		$x_{n-1T-1k}$
n	T	y_{nT}	x_{nT1}	x_{nT2}	...	x_{nTk}

Zdroj: vlastní zpracování

Panelová data mohou být zobrazována v různých formátech. Některé zdroje rozlišují široký (wide) a dlouhý (long) formát. **Dlouhý formát** shromažďuje pozorování každé proměnné ze všech skupin ve všech časových obdobích do jednoho sloupce. Naproti tomu **široký formát** zobrazuje opakované pozorování objektu na jednom řádku a každou proměnnou v samostatném sloupci.¹⁰

Tyto dva rozdílné formáty se využívají v závislosti na tom, jaká je použita analýza panelových dat a v jakém softwaru jsou data zpracovávána.¹¹ Pro regresní analýzu je spíše využíván dlouhý formát. Ten je také považován za efektivnější, vzhledem k tomu, že pokud se v širokém formátu zvýší počet proměnných a časových okamžiků, může být datová matice až příliš velká.¹²

V softwaru Gretl, ve kterém budou zpracována data v praktické části této práce, se pracuje s dlouhým formátem. V Tabulka 3 a Tabulka 4 jsou datové formáty zobrazeny.

Tabulka 3 Dlouhý formát panelových dat

objekt	čas	závislá proměnná
i	t	y_{it}
1	1	y_{11}
1	2	y_{12}
1	T	y_{1T}
...
n	1	y_{n1}
n	2	y_{n1}
n	T	y_{nT}

Zdroj: vlastní zpracování dle Apteck, 2019

¹⁰ *Introduction to the Fundamentals of Panel Data*. Online. Apteck. 2019. Dostupné z: <https://www.aptech.com/blog/introduction-to-the-fundamentals-of-panel-data/>. [cit. 2024-06-02]

¹¹ *The wide and long data format for repeated measures data*. Online. The Analysis Factor. b.r. Dostupné z: <https://www.theanalysisfactor.com/wide-and-long-data/>. [cit. 2024-06-17].

¹² ANDREß, Hans-Jurgen; GOLDSCH, Katrin a W.SCHMIDT, Alexander. *Applied Panel Data Analysis for Economic and Social Surveys*. Springer, 2013. s. 17. ISBN 987-3-642-32913-5.

Tabulka 4 Široký formát panelových dat

čas	závislá proměnná		
t	y_1	y_2	y_n
1	y_{11}	y_{21}	y_{n1}
2	y_{12}	y_{22}	y_{n2}
3	y_{13}	y_{23}	y_{n3}
4	y_{14}	y_{24}	y_{n4}
...
T-1	y_{1T-1}	y_{2T-1}	y_{nT-1}
T	y_{1T}	y_{2T}	y_{nT}

Zdroj: vlastní zpracování dle Apteck, 2019

Panelová data lze dále členit podle panelové struktury na:

- krátké panelové soubory – kdy je sledováno mnoho objektů jen v několika málo časových okamžicích,
- dlouhé panelové soubory – kdy je sledováno několik objektů ve výrazně vyšším počtu časových okamžiků.¹³

Krátké panelové soubory jsou typické pro mikroekonomické výzkumy, kdy je zkoumáno mnoho jednotlivců během několika let. Naopak dlouhé panelové soubory se častěji vyskytují u makroekonomických výzkumů, kdy se zkoumají dlouhé časové řady makroekonomických veličin několika států. Může se jednat například o výzkumy států Evropské Unie nebo států skupiny G7 v rámci několika desítek let.¹⁴

¹³ GUJARATI, Damodar N. a PORTER, Dawn C. *Basic econometrics*. 5th ed. McGraw-Hill/Irwin, 2009. s. 593. ISBN 978-0-07-337577-9.

¹⁴ BALTAGI, Badi H. *Econometric Analysis of Panel Data*. 6th ed. Cham: Springer Cham, 2021. s. 1. ISBN 9783030539528.

1.2.1 Panel

Panel je skupina objektů se stejnými nebo velmi podobnými vlastnostmi, která je kontinuálně zkoumána. Objekty mohou představovat například osoby, domácnosti, firmy nebo státy.¹⁵

Panely můžeme rozlišovat na **vyrovnané** (vyvážené) a **nevyrovnané** (nevyvážené).

V ideálním případě žádné hodnoty nechybí a jedná se o vyrovnaný panel. Pokud tedy panel obsahuje N objektů a T období, pro počet pozorování n musí platit, že $n = N \cdot T$. Vyrovnaný panel může mít několik podob. Nezáleží na tom, zda se počet objektů rovná počtu časových období.

Panel je nevyrovnaný, pokud je počet pozorování menší než $N \cdot T$.¹⁶

Nevyrovnaný panel může mít několik následujících podob:

- různé počáteční časové okamžiky sledování jednotlivých objektů,
- různé konečné časové okamžiky sledování jednotlivých objektů,
- různé počáteční i konečné časové okamžiky sledování jednotlivých objektů,
- chybějící hodnoty napříč panelem.

Pokud lze najít časový úsek, pro který jsou všechna pozorování zaznamenána, je možné jej zkrátit. Lze tak převést panel na vyrovnaný. Dalším příkladem nevyrovnaného panelu je takový, kde chybí jen některé hodnoty. Tyto hodnoty lze zpětně odhadnout. Při odhadu vyššího počtu hodnot je ale potřeba počítat s možností zkreslených výsledků.¹⁶

Některé modely panelových dat jsou určeny pouze pro vyrovnané panely. Někdy je nutné zkrátit nevyrovnané panely tak, aby obsahovaly období, pro která existují všechna pozorování nebo odhadnout chybějící hodnoty.¹⁷

V rámci některých výzkumů není možné sledovat stále stejný panel. Jedná se například o průzkumy spotřebitelského chování v rozvíjejících se zemích, kdy není možné sledovat

¹⁵ TAHAL, Radek. *Marketingový výzkum: postupy, metody, trendy*. 2. vydání. *Expert (Grada)*. Praha: Grada Publishing, 2022. s. 71. ISBN 978-80-271-3535-6.

¹⁶ *Detailed Explanation of Panel Data*. Online. Medium. 2022. Dostupné z: <https://medium.com/geekculture/detailed-explanation-of-panel-data-how-to-identify-balanced-and-unbalanced-panel-data-fd973fa788ae>. [cit. 2024-06-02].

¹⁷ *Introduction to the Fundamentals of Panel Data*. Online. Aptech. 2019. Dostupné z: <https://www.aptech.com/blog/introduction-to-the-fundamentals-of-panel-data/>. [cit. 2024-06-02].

v průběhu let stále stejné jednotlivce. Řešením je sdružování jednotlivců do kohort na základě určitých vlastností pozorovaných u všech jednotlivců, které se v průběhu času nemění (například rok narození) a vytvoření tzv. **pseudopanelu**. Při jeho analýze se pak nesledují jednotlivci, ale jednotlivé kohorty.¹⁸

1.3 Výhody a problémy panelových dat

V následujícím textu je uvedeno několik výhod a problémů, které panelová data, na rozdíl od dat časových nebo průřezových, přinášejí.

Výhody

Mezi výhody patří velké množství pozorování. To zapříčiňuje zvýšení stupňů volnosti a snížení kolinearity mezi vysvětlujícími proměnnými. Tím dochází ke zpřesnění odhadů parametrů v modelu. Některé ekonomické otázky tak lze zodpovědět pouze pomocí analýzy panelových dat, nikoliv pouze analýzou dat časových nebo průřezových.¹⁹

Analýza panelových dat umožňuje kontrolovat individuální heterogenitu. Při použití analýzy časových nebo průřezových dat nedochází k této kontrole, a tudíž může docházet ke zkreslení výsledků. Mezi další výhody patří schopnost studovat dynamiku změn, které nejsou na jednotlivých průřezech patrné. Teprve až průřezy sledované v čase dokážou změny odhalit.²⁰

Analýza panelových dat je také schopna lépe identifikovat a měřit efekty, které mohou na data působit. Lze tak lépe studovat chování jednotlivých datových souborů. Baltagi²⁰ uvádí následující příklad. Existuje průřez žen s 50% průměrnou roční mírou účasti na trhu práce. Pouze panelová data dokážou rozlišit případ, kdy je každá žena zaměstnána z 50 % od případu, kdy je 50 % žen zaměstnáno a zbytek je nezaměstnaných.

Hsiao²¹ jako další výhodu uvádí, že analýza panelových dat může přinášet přesnější předpovědi chování jedinců díky sdružování dat. Odhadovat parametry lze i v případě, že jsou data o

¹⁸ BALTAGI, Badi H. *Econometric Analysis of Panel Data*. 6th ed. Cham: Springer Cham, 2021. s. 265. ISBN 9783030539528.

¹⁹ HSIAO, Cheng. *Analysis of Panel Data*. 4th ed. Cambridge University Press, 2022. s. 3. ISBN 978-1-316-51210-4.

²⁰ BALTAGI, Badi H. *Econometric Analysis of Panel Data*. 6th ed. Cham: Springer Cham, 2021. s. 6. ISBN 9783030539528.

²¹ HSIAO, Cheng. *Analysis of Panel Data*. 4th ed. Cambridge University Press, 2022. 7–8. ISBN 978-1-316-51210-4.

chování jednotlivce omezená. Předpověď totiž nevychází pouze z chování daného jednotlivce, ale bere v úvahu chování všech jedinců v panelu.

Mezi další výhodu patří snížení statistické chyby. Díky sledování změn na stále stejném panelu dochází k eliminování výběrové chyby. Nepochází tak k náhodnému kolísání časových řad.²²

Nevýhody

Hsiao²¹ řadí mezi nevýhody panelových dat nepozorovanou heterogenitu. Panelová data jsou získávána dlouhodobým sledováním, během kterého na objekt (např. jedinec, domácnost) působí velké množství efektů. Ty mají vliv na chování závisle proměnné. Tyto efekty nelze a ani není žádoucí všechny zahrnovat do modelu. Faktory, které nemají na výsledné odhady významný vliv, se v modelu vynechávají.

Baltagi²³ uvádí, že ke zkresleným výsledkům může docházet, pokud je v panelu sledováno několik objektů ve výrazně větším počtu časových okamžiků. Existuje tedy několik málo dlouhých časových řad. Nebo naopak mají panelová data takovou strukturu, kdy je mnoho objektů sledováno jen v několika málo časových okamžicích.

Dalším problémem, který se může vyskytnout při realizaci panelového výzkumu, patří změny ve zkoumaném panelu. V praxi není téměř možné, aby panel zůstal dlouhodobě bez jakékoliv obměny, především při rozsáhlých výzkumech, kdy se data sbírají například několik let. Při analýze dat je tedy potřeba pracovat s tím, že některá data mohou chybět. Panel se průběžně obměňuje v závislosti na tom, jak respondenti ukončují spolupráci na panelovém výzkumu. Děje se tak z různých důvodů na straně respondenta i zpracovatele dat.

Mezi důvody obměny panelu se řadí například:

- změna životní situace respondenta, kdy respondent již nesplňuje podmínky účasti v panelu (například zestárne),
- vlastní rozhodnutí respondenta,
- ukončení spolupráce s respondentem z důvodu neplnění povinností.

²² TAHAL, Radek. *Marketingový výzkum: postupy, metody, trendy*. 2. vydání. *Expert (Grada)*. Praha: Grada Publishing, 2022. s. 72. ISBN 978-80-271-3535-6.

²³ BALTAGI, Badi H. *Econometric Analysis of Panel Data*. 6th ed. Cham: Springer Cham, 2021. s. 11. ISBN 9783030539528.

Obměna musí proběhnout tak, aby panel zůstal reprezentativní.²⁴

1.4 Testy jednotkových kořenů

Testy jednotkových kořenů souvisí s předpokladem stacionarity. Modely panelových dat předpokládají stacionární procesy. Nicméně většina ekonomických procesů je nestacionární. Pokud je časová řada nestacionární, obsahuje jednotkový kořen. Arlt definuje stacionární proces jako: „*Stochastický proces, jehož charakteristiky náhodných veličin jsou v čase neměnné*“.²⁵

Testy jednotkových kořenů panelových dat vycházejí z testů jednotkových kořenů časových řad. Testování jednotkových kořenů panelových dat s sebou přináší, na rozdíl od testování jednotkových kořenů časových řad, následující problémy:

- modely zahrnují velké množství nepozorované heterogenity,
- v některých případech se nesprávně předpokládá nezávislost průřezových dat,
- v případě zamítnutí nulové hypotézy o neexistenci jednotkových kořenů je složité interpretovat výsledky panelového modelu.

Z těchto důvodů byly vytvořeny testy jednotkových kořenů, které jsou specifické pro panelová data, nicméně stále vycházejí z testů pro časové řady.

Jejich vývoj začal v 90. letech 20. století (Quah – 1992, 1994).²⁶ Následovali další autoři.

Baltagi²⁷ uvádí testy jednotkových kořenů od následujících autorů:

- Levin, Lin a Chu test,
- Im, Pesaran a Shin test,
- Breitung test.

²⁴ TAHAL, Radek. *Marketingový výzkum: postupy, metody, trendy*. 2. vydání. *Expert (Grada)*. Praha: Grada Publishing, 2022. s. 80–81. ISBN 978-80-271-3535-6.

²⁵ ARLT, Josef a ARLTOVÁ, Markéta. *Ekonomické časové řady*. Praha: Professional Publishing, 2009. s. 25. ISBN 978-80-86946-85-6.

²⁶ KOČENDA, Evžen a ČERNÝ, Alexandr. *Elements of time series econometrics: an applied approach*. Third edition. Prague: Charles University in Prague, Karolinum Press, 2015. ISBN 978-80-246-3199-8.

²⁷ BALTAGI, Badi H. *Econometric Analysis of Panel Data*. 6th ed. Cham: Springer Cham, 2021. s. 187. ISBN 9783030539528.

1.5 Oblasti využití panelových dat

Analýza panelových dat nachází využití v nejrůznějších oblastech. V rámci podnikové praxe pomáhá při finanční analýze podniku. Cenným nástrojem jsou panelová data také v oblasti marketingu. Pomocí analýzy panelových dat může firma lépe pochopit chování spotřebitelů nebo sledovat, jak funguje jejich marketingová komunikace. A přizpůsobit jí tak, aby byla co nejvíce účinná. Díky panelovým datům je také firma schopná lépe pochopit fungování trhu a následně pak přizpůsobit své produkty nebo strategie.²⁸

Analýza panelových dat se využívá i v oblasti makroekonomie. Slouží pro analýzu makroekonomických ukazatelů v rámci více zemí. Dále se panelová regrese využívá například ke zkoumání vlivu ekonomického růstu země na rovnost mužů a žen ve společnosti.²⁹

Další významnou oblastí, kde se analýza panelových dat využívá, je sociologie. Zde slouží ke studiu změn chování jednotlivců nebo domácností a pomáhá určit faktory, které stojí za určitými sociologickými jevy. Příkladem může být polský sociologický panelový výzkum POPAN, který byl zahájen v roce 1988 a zabývá se mimo jiné sociální strukturou obyvatel, nerovností ve společnosti nebo emigrací.³⁰

Analýzou panelových dat lze také studovat zdravotní stav pacientů. Na základě dlouhodobého sledování jejich zdravotního stavu je možné lépe stanovovat prevenci a léčbu.³¹

1.6 Panelový výzkum

Následující podkapitola se zaměřuje především na panelové šetření v rámci podnikové praxe, tedy například marketingové panelové šetření nebo šetření spotřebitelského chování.

Panelový výzkum je využíván v případě, kdy je cílem zachytit vývoj stejné skupiny respondentů v čase. Pokud není potřeba studovat stále stejnou skupinu respondentů, je využíván

²⁸ *Inform your marketing strategies with panel data.* Online. Kantar. 2023. Dostupné z: <https://www.kantar.com/north-america/inspiration/research-services/inform-your-marketing-strategies-with-panel-data-pf>. [cit. 2024-06-23].

²⁹ *Jak na GPG z pohledu makroekonomie.* Online. Ministerstvo práce a sociálních věcí, 2020. ISBN 978-80-7421-239-0. Dostupné z: https://www.rovnaodmena.cz/www/img/uploads/Makroanalzya_07_22022021.pdf. [cit. 2024-06-23].

³⁰ *About POLPAN.* Online. POLPAN. B.r. Dostupné z: <https://polpan.org/en/about-polpan/>. [cit. 2024-06-23].

³¹ *The Role of Panel Research in Enhancing Patient Experience in the Healthcare Industry.* Online. Data Spring. 2023. Dostupné z: <https://www.d8aspring.com/blog/the-role-of-panel-research-in-the-healthcare-industry>. [cit. 2024-06-23].

klasický výzkum, který je opakován v čase. Ten se používá například pro sledování vývoje poslechovosti rádií.³²

Nejdříve je potřeba sestavit panel respondentů tak, aby byl reprezentativní. Panel respondentů se tvoří s ohledem na předem definované cíle výzkumu. Panely mohou být koncipované jako dlouhodobé nebo krátkodobé. Krátkodobé panely se vyznačují rotací respondentů. Respondenti se v takovém panelu mění cíleně a řízeně, panelové šetření má v tomto případě více vln. Intervaly mezi vlnami jednotlivých výzkumů se mohou značně lišit. Některá šetření mohou probíhat jednou měsíčně, jiné jednou za několik let. Krátkodobé panely jsou vhodné pro výzkumy, kde je vyžadována vysoká reprezentativnost populace. U dlouhodobého panelu se vyskytuje problém s dostatečnou motivací respondenta zůstat v panelu.³³

Panelové šetření může probíhat několika způsoby. Data se mohou získávat prostřednictvím kontinuálního dotazníkového šetření na stále stejné skupině respondentů nebo z deníků, kam respondenti pravidelně vyplňují požadované údaje. Lze použít i retrospektivní otázky a vyhnout se tak panelovému šetření. Výsledky výzkumu pak ale nemusí být přesné. Dále mohou být data získávána také z elektronických měřičů nebo aplikací určených pro sledování dat. Tento způsob je využíván například pro měření venkovní reklamy nebo sledovanosti TV. Pomocí panelového šetření se také zkoumá nákupní chování spotřebitelů. Domácnosti, zapojené do tohoto šetření, pomocí speciální čtečky zaznamenávají čárové kódy zboží, které nakupují.³⁴

Při realizaci panelového výzkumu lze zvolit mezi dvěma způsoby sběru dat:

- data jsou zaznamenávány v reálném čase,
- data jsou zaznamenány retrospektivně.

Výběr záleží na cíli a charakteru výzkumu. Retrospektivní sběr dat není tak časově ani finančně náročný jako kontinuální zaznamenávání dat. Pro některé výzkumy, ale nemusí být dostupná data dostačující.

Výhodou využití panelového výzkumu může být úspora nákladů, které vznikají v souvislosti s hledáním a zaškolováním respondentů. Čím frekventovanější dotazování je, tím výhodnější

³² TAHAL, Radek. *Marketingový výzkum: postupy, metody, trendy*. 2. vydání. *Expert (Grada)*. Praha: Grada Publishing, 2022. s. 71–72. ISBN 978-80-271-3535-6.

³³ LIPPS, Oliver. Panel Surveys. Online. *FORS Guide*. Č. no.14. Dostupné z: https://forscenter.ch/wp-content/uploads/2021/03/forsguide_panel_surveys_lipps_v_1.pdf. [cit. 2024-06-25].

³⁴ TAHAL, Radek. *Marketingový výzkum: postupy, metody, trendy*. 2. vydání. *Expert (Grada)*. Praha: Grada Publishing, 2022. s. 73–74. ISBN 978-80-271-3535-6.

je mít stále stejnou skupinu respondentů. Příkladem je analýza televizní sledovanosti, kde se data získávají každý den a náklady na získání respondenta jsou vysoké. V tomto případě by bylo možné realizovat klasický kontinuální výzkum, ale levnější je realizace panelového výzkumu.³⁵

Mezi nevýhody panelového šetření patří nízká míra retence respondentů. V průběhu šetření, které je prováděno dlouhodobě, dochází k takzvanému „opotrebení“ respondentů, kteří panel opustí. Pokud je opotrebení selektivní, může docházet ke zkresleným odhadům v panelovém modelu.³⁶

Problémem u panelových šetření může být takzvaná profesionalizace respondentů. Děje se tak z důvodu dlouhodobého charakteru výzkumu. Někteří respondenti nemusí být poctiví při pravidelném odpovídání na otázky. Vzhledem k tomu, že vyplňují dotazník pravidelně, nemusí se nad odpověďmi příliš zamýšlet. Někteří respondenti mohou mít i potřebu záměrně zkreslovat odpovědi, aby například podpořili svojí oblíbenou značku. Tyto změny chování se označují jako panelový efekt, ten zkresluje výsledky panelového šetření. Se spoluprací na panelovém výzkumu bývá spojena finanční odměna, ta by měla být nastavena tak, aby neovlivňovala chování respondentů v průběhu výzkumu.³⁷

1.6.1 Příklady panelových výzkumů

Hsiao³⁸ uvádí, že v USA patří k nejdéle trvajícím panelovým výzkumům National Longitudinal Surveys of Labor Market Experience (NLS) a Panel Study of Income Dynamics (PSID). První ze zmiňovaných je soubor průzkumů, který shromažďuje data o trhu práce. Průzkum začal již v roce 1966. Účastníci jsou rozděleny do skupin podle věku a je sledováno mnoho jejich charakteristik jako například jejich aktivity na trhu práce, celoživotní vzdělávání, plodnost, zdravotní nebo rodinný stav.³⁹

³⁵ TAHAL, Radek. *Marketingový výzkum: postupy, metody, trendy*. 2. vydání. *Expert (Grada)*. Praha: Grada Publishing, 2022. s. 71–74. ISBN 978-80-271-3535-6.

³⁶ LIPPS, Oliver. Panel Surveys. Online. *FORS Guide*. Č. 14. Dostupné z: https://forscenter.ch/wp-content/uploads/2021/03/forsguide_panel_surveys_lipps_v_1.pdf. [cit. 2024-06-25].

³⁷ TAHAL, Radek. *Marketingový výzkum: postupy, metody, trendy*. 2. vydání. *Expert (Grada)*. Praha: Grada Publishing, 2022. s. 83. ISBN 978-80-271-3535-6.

³⁸ HSIAO, Cheng. *Analysis of Panel Data*. 4th ed. Cambridge University Press, 2022. s. 1. ISBN 978-1-316-51210-4.

³⁹ *Cohorts*. Online. National Longitudinal Surveys. b.r. Dostupné z: <https://www.nlsinfo.org/content/cohorts>. [cit. 2024-05-23].

Panelový výzkum PSID začal v roce 1968. Jedná se o nejdéle probíhající průzkum domácností, který je řízen Michiganskou univerzitou. Je do něj zahrnuto přes 18 000 obyvatel USA, kteří žijí v přibližně pěti tisících domácnostech. Shromažďují se data o zaměstnání, příjmech a výdajích, zdravotním stavu, nebo vzdělání. Sběr dat probíhá pomocí rozhovoru, který poskytne jedna osoba z domácnosti. Do roku 1997 tyto rozhovory probíhaly každoročně, od té doby pouze jednou za dva roky.⁴⁰

V České republice probíhá od roku 2015 panelové šetření domácností CHPS. Jedná se doposud o jediný panelový výzkum domácností v České republice.⁴¹ Šetření probíhalo v letech 2015 až 2019. V prvním roce se ho účastnilo přes pět tisíc domácností, poté počet domácností klesal. Kvůli pandemii šetření pokračovalo až v roce 2023, kdy se zapojilo přes tři tisíce domácností. Na rozdíl od amerického výzkumu PSID je dotazováno více členů z jedné domácnosti, to umožňuje zkoumat i vývoj rodinných vazeb. Dotazování probíhá kombinací osobních rozhovorů a vyplňování dotazníku přes internet.⁴²

Tento výzkum poskytuje informace o stavu české společnosti. Analýzou získaných dat je možné zjistit hospodářskou a ekonomickou situaci, chování na trhu práce nebo informace ohledně trávení volného času nebo zdraví. Výsledky šetření poskytují i informace o zájmu o politické dění.⁴³

V rámci Organizace pro hospodářskou spolupráci a rozvoj je realizován mezinárodní výzkum PIAAC (Programme for the International Assessment of Adult Competencies). Jedná se o výzkum, který hodnotí dovednosti u dospělých. Je dotazováno zhruba 5 000 dospělých v každé ze 40 zúčastněných zemí. Hodnotí se jejich literární a matematická gramotnost. Již se uskutečnily dva cykly výzkumu. První proběhl mezi lety 2011–2018, proběhly tři kola sběru dat. Sběr dat v rámci druhého cyklu začal v roce 2018. Další kola by se měla konat v letech 2024–2029.⁴⁴

⁴⁰ *Panel Study of Income Dynamics*. Online. PSID. c2024. Dostupné z: <https://psidonline.isr.umich.edu>. [cit. 2024-05-23].

⁴¹ *Představení výzkumu*. Online. CHPS. c2023. Dostupné z: <https://chps.soc.cas.cz/cs/predstaveni-vyzkumu>. [cit. 2024-05-23].

⁴² *Dotazování v 6. ročníku bylo úspěšně ukončeno!*. Online. CHPS. 2024. Dostupné z: <https://chps.soc.cas.cz/cs/aktuality/dotazovani-v-6-rocniku-bylo-uspesne-ukonceno>. [cit. 2024-05-23].

⁴³ *Představení výzkumu*. Online. CHPS. c2023. Dostupné z: <https://chps.soc.cas.cz/cs/predstaveni-vyzkumu>. [cit. 2024-05-23].

⁴⁴ *The Survey of Adult Skills (PIAAC)*. Online. OECD. b.r. Dostupné z: <https://www.oecd.org/skills/piaac/about/#d.en.481111>. [cit. 2024-05-29].

Na tento výzkum navazují různá panelová šetření. V České republice například v návaznosti na výzkum PIAAC proběhlo panelové šetření SKILLS, které sledovalo, jak závisí změny podmínek na trhu práce na získaných dovednostech jednotlivce. Mimo jiného se také sledovala profesní dráha účastníků výzkumu. Bylo zjišťováno, jak dokážou skloubit profesní a osobní život nebo se zkoumaly rozdíly v rovném odměňování mužů a žen.⁴⁵

⁴⁵ *Národní výzkumy*. Online. PIAAC. c2011–2024. Dostupné z: <https://piaac.cz/o-vyzkumu/navazne-vyzkumy/>. [cit. 2024-05-29].

2. KONSTRUKCE VYBRANÝCH MODELŮ PANELOVÝCH DAT

V následující kapitole jsou rozebrány vybrané modely panelových dat. Konkrétně se jedná o spojený regresní model. Dále jsou popsány dva základní modely individuálních efektů, a to model s fixními efekty a model s náhodnými efekty. Následně je popsán dynamický model panelových dat a jeho speciální varianta – autoregresivní model panelových dat.

2.1 Panelová regrese

Regresní analýza označuje statistické metody, za jejichž pomoci se data modelují. Slouží pro předpovědi hodnot jedné náhodné veličiny v závislosti na hodnotách druhé náhodné veličiny.⁴⁶ Panelová regrese zkoumá vztahy mezi proměnnými v rámci průřezů i v rámci časových řad.⁴⁷

Modely panelové regrese zahrnují závisle proměnnou, jednu nebo více nezávislých proměnných. Dále náhodné složky, které vyjadřují náhodné veličiny působící na závislou proměnnou a nelze v modelu vyjádřit.⁴⁸

Základní regresní model panelových dat lze vyjádřit následovně:

$$y_{it} = \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + \alpha_1 Z_{i1} + \alpha_2 Z_{i2} + \dots + \alpha_q Z_{iq} + u_{it} \quad (2.1)$$

kde:

y_{it} je vysvětlovaná proměnná jednotky i ($i = 1, 2, \dots, n$) a čase t ($t = 1, 2, \dots, T$),

β je odhadovaný parametr,

x_1 až x_k jsou vysvětlující proměnné,

z_1 až z_q jsou proměnné, které představují individuální efekty,

u_{it} je reziduální složka pro jednotku i v časovém okamžiku t .

⁴⁶ KUBANOVÁ, Jana. *Statistické metody pro ekonomickou a technickou praxi*. 3. dopl. vyd. Statisticky, 2008. s. 105. ISBN 978-80-85659-47-4.

⁴⁷ *Regression using Panel data for beginners*. Online. Medium. 2023. Dostupné z: <https://medium.com/@laibamehnaz/regression-using-panel-data-for-beginners-f9b636deb459>. [cit. 2024-06-23].

⁴⁸ WOOLDRIDGE, Jeffrey M. *Introductory Econometrics: A Modern Approach*. 3rd ed. Mason: Thomson South-Western, 2006. s. 25. ISBN 0-324-32348-4.

2.2 Spojený regresní model

Jako základní model v oblasti panelových dat je považován spojený regresní model, v literatuře někdy nazývaný jako poolový (z angl. pooled model). Tento model využívá metodu nejmenších čtverců (OLS – Ordinary Least Squares). Metoda OLS ignoruje, že se jedná o panelová data. Všechna data jsou seřazena do jedné dlouhé časové řady. Spojený regresní model nepředpokládá, že existují individuální efekty rozdílné pro jednotlivé proměnné. V tomto případě jsou tedy všechny individuální efekty zahrnuté do jedné konstanty, která pro jednotlivé proměnné zůstává stále stejná.⁴⁹

Spojený regresní model lze vyjádřit následovně:

$$y_{it} = \alpha + \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + u_{it} \quad (2.2)$$

kde:

- y_{it} je vysvětlovaná proměnná jednotky i ($i = 1, 2, \dots, n$) a čase t ($t = 1, 2, \dots, T$),
- α vyjadřuje všechny individuální efekty,
- β vyjadřuje odhadovaný parametr,
- x_{itk} je hodnota j -té vysvětlující proměnné ($j = 1, 2, \dots, k$) pro i -tou jednotku v časovém okamžiku t ,
- u_{it} je reziduální složka pro jednotku i v časovém okamžiku t .

Wooldridge⁵⁰ uvádí několik následujících podmínek pro použití spojeného regresního modelu:

- předpoklad náhodného výběru,
- každá vysvětlující proměnná se v průběhu času mění (alespoň u některých objektů) a mezi vysvětlujícími proměnnými neexistují žádné dokonalé lineární vztahy,
- neexistují nepozorované individuální efekty, které ovlivňují závislou proměnnou,
- rozptyl reziduální složky je pro všechny vysvětlující proměnné konstantní,
- reziduální složky jsou nekorelované.

⁴⁹ *The Pooled OLS Model*. Online. Tilburg Science Hub. B.r. Dostupné z: <https://tilburgsciencehub.com/topics/analyze/causal-inference/panel-data/pooled-ols/>. [cit. 2024-06-23].

⁵⁰ WOOLDRIDGE, Jeffrey M. *Introductory Econometrics: A Modern Approach*. 3rd ed. Mason: Thomson South-Western, 2006. s. 483. ISBN 0-324-32348-4.

2.3 Modely s individuálními efekty

Modely individuálních efektů řeší problém heterogenity v panelových datech. Individuální efekty představují různorodost, kterou se od sebe jednotlivé objekty mohou odlišovat a během času se nemění. Green⁵¹ mezi individuální efekty řadí pohlaví, bydliště, rodinné zázemí, zdravotní stav, vzdělání, zaměstnání nebo ekonomickou a politickou situaci země.

Problém individuálních efektů řeší dva typy modelů, a to model s fixními efekty a model s náhodnými efekty.

Gujarati⁵² ve své práci uvádí několik pravidel, které pomáhají rozhodnout, zda použít model s fixními nebo náhodnými efekty. Mezi ně patří:

- pokud je počet časových okamžiků vysoký a počet objektů nízký, hodnoty odhadnutých parametrů budou podobné, v tom případě se preferuje model fixních efektů,
- pokud je naopak počet sledovaných objektů vysoký a počet časových okamžiků nízký, rozdíly mezi odhadnutými parametry budou velké, pokud jsou průřezové jednotky náhodně vybrané, preferuje se model náhodných efektů (jinak opět model fixních efektů).

Tyto pravidla nemusí dostatečně přesně určit, jaký model použít. Pro přesnější určení se používá **Hausmannův test**.⁵³ Ten určuje, zda je odhad parametrů s fixními a náhodnými efekty výrazně odlišný. Testuje nulovou hypotézu:

H_0 : náhodné efekty jsou konzistentní a efektivní,

oproti alternativní:

H_1 : náhodné efekty jsou nekonzistentní.

V případě zamítnutí nulové hypotézy je vhodné použít metodu fixních efektů, protože náhodná proměnná je pravděpodobně korelována s jednou nebo více nezávislými proměnnými. V případě, že není zamítnuta nulová hypotéza, je možné použití obou metod.⁵⁴

⁵¹ GREENE, William H. *Econometric Analysis*. 5th ed. Pearson Education, 2003. s. 285. ISBN 0-13-066189-9.

⁵² GUJARATI, Damodar N. a PORTER, Dawn C. *Basic econometrics*. 5th ed. McGraw-Hill/Irwin, 2009. s. 606–607. ISBN 978-0-07-337577-9.

⁵³ WOOLDRIDGE, Jeffrey M. *Introductory Econometrics: A Modern Approach*. 3rd ed. Mason: Thomson South-Western, 2006. s. 497–498. ISBN 0-324-32348-4.

⁵⁴ BALTAGI, Badi H. *Econometric Analysis of Panel Data*. 6th ed. Cham: Springer Cham, 2021. s. 11. ISBN 9783030539528.

Další možností, kterou lze využít při rozhodování mezi modelem s fixními nebo náhodnými efekty je **Breusch-Paganův test**, který testuje nulovou hypotézu:

H_0 : rozptyl náhodných složek je nulový,

oproti alternativní:

H_1 : rozptyl náhodných složek není nulový.

V případě zamítnutí nulové hypotézy je vhodné použít metodu náhodných efektů.⁵⁵

2.3.1 Model s fixními efekty

Panelový model s fixními efekty (angl. Fixed Effects Model – FEM) předpokládá, že individuální efekty jsou nepozorované, ale korelované s vysvětlujícími proměnnými. Rozdílnost mezi průřezovými jednotkami tedy určuje pouze parametr α_i , který absorbuje všechny nepozorované faktory.

Parametr je na rozdíl od spojeného regresního modelu specifický pro každou průřezovou jednotku. S rostoucím počtem sledovaných objektů roste i počet parametrů α_i , které je potřeba v modelu odhadnout.

Lze vyjádřit takto⁵⁶:

$$y_{it} = \alpha_i + \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + u_{it} \quad (2.3)$$

$$\alpha_i = \alpha_1 z_{i1} + \alpha_2 z_{i2} + \dots + \alpha_q z_{iq} \quad (2.4)$$

kde:

- y_{it} je vysvětlovaná proměnná jednotky i ($i = 1, 2, \dots, n$) a čase t ($t = 1, 2, \dots, T$),
- α_i vyjadřuje individuální efekty z_1 až z_q ,
- β vyjadřuje odhadovaný parametr,
- x_{itk} je hodnota j -té vysvětlující proměnné ($j = 1, 2, \dots, k$) pro i -tou jednotku v časovém okamžiku t ,
- u_{it} je reziduální složka pro jednotku i v časovém okamžiku t .

⁵⁵ GUJARATI, Damodar N. a PORTER, Dawn C. *Basic econometrics*. 5th ed. McGraw-Hill/Irwin, 2009. s. 605. ISBN 978-0-07-337577-9.

⁵⁶ WOOLDRIDGE, Jeffrey M. *Introductory Econometrics: A Modern Approach*. 3rd ed. Mason: Thomson South-Western, 2006. s. 486. ISBN 0-324-32348-4.

Metodu fixních efektů lze bez problému použít na vyvážené panely. U nevyvážených panelů je nutné předpokládat, že důvod chybějících časových hodnot není systematicky spojen s reziduální složkou. Pokud například jednotlivec opustí zkoumaný vzorek z důvodu, který koreluje s nepozorovanými faktory, které se v průběhu času mění a ovlivňují závislou proměnnou, může být odhad zkreslený.⁵⁷

Do modelu s fixními efekty lze zahrnout pomocné proměnné, které nabývají binárních hodnot. Lze tak rozlišit některé proměnné na dvě skupiny.

2.3.2 Model s náhodnými efekty

Panelový model s náhodnými efekty (angl. Random Effects Model – REM) předpokládá, že individuální efekty jsou nepozorované, ale nekorelované s vysvětlujícími proměnnými. Proto se k parametru individuálních efektů α přičítá náhodná složka, která je pro jednotlivé průřezové jednotky rozdílná.⁵⁸

Model lze vyjádřit takto:

$$y_{it} = \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + (\alpha + \varepsilon_i) + u_{it} \quad (2.5)$$

kde:

- y_{it} je vysvětlovaná proměnná jednotky i ($i = 1, 2, \dots, n$) a čase t ($t = 1, 2, \dots, T$),
- β vyjadřuje odhadovaný parametr,
- x_{itk} je hodnota j -té vysvětlující proměnné ($j = 1, 2, \dots, k$) pro i -tou jednotku v časovém okamžiku t ,
- α vyjadřuje individuální efekty pro všechny jednotky i stejné,
- ε_i je reziduální složka pro jednotku i ,
- u_{it} je reziduální složka pro jednotku i v časovém okamžiku t .

Model náhodných efektů tedy rozlišuje nepozorovanou heterogenitu v rámci jednotlivých objektů (konstantní v čase) a heterogenitu v rámci času.⁵⁹

⁵⁷ WOOLDRIDGE, Jeffrey M. *Introductory Econometrics: A Modern Approach*. 3rd ed. Mason: Thomson South-Western, 2006. s. 493. ISBN 0-324-32348-4.

⁵⁸ WOOLDRIDGE, Jeffrey M. *Introductory Econometrics: A Modern Approach*. 3rd ed. Mason: Thomson South-Western, 2006. s. 486. ISBN 0-324-32348-4.

⁵⁹ *The Random Effects Model*. Online. Tilburg Science Hub. B.r. Dostupné z: <https://tilburgsciencehub.com/topics/analyze/causal-inference/panel-data/random-effects/>. [cit. 2024-06-23].

2.4 Dynamické modely

Předchozí zmiňované modely se řadí mezi statické. Statické modely modelují pouze současné vztahy, nedovolují ale zkoumat vliv chování předchozích hodnot na současné hodnoty. Tento typ modelů se využívá pro dynamické procesy, ve kterých jsou proměnné závislé na své předchozí hodnotě. Dynamické procesy jsou typické pro ekonomickou oblast.⁶⁰ Důvodem jsou neustále vyvíjející se ekonomické podmínky. U těchto procesů je velmi obtížné předpovídat následující hodnoty.

Do dynamických modelů panelových dat se zahrnuje zpožděná závisle proměnná, která je korelovaná s reziduální složkou.⁶¹ Vzhledem k tomu není vhodné pro odhady používat metodu nejmenších čtverců.

Dynamický model panelových dat lze vyjádřit takto:

$$y_{it} = \gamma y_{t-1} + \alpha_i + \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + \lambda_t + u_{it} \quad (2.6)$$

kde:

- y_{it} je vysvětlovaná proměnná jednotky i ($i = 1, 2, \dots, n$) a čase t ($t = 1, 2, \dots, T$),
- y_{t-1} je zpožděná závisle proměnnou,
- α_i jsou individuální efekty rozdílné pro každé i ,
- β je odhadovaný parametr,
- x_{itk} je hodnota j -té vysvětlující proměnné ($j = 1, 2, \dots, k$) pro i -tou jednotku v časovém okamžiku t ,
- λ_t jsou časově specifické efekty,
- u_{it} je reziduální složka pro jednotku i v časovém okamžiku t .

Parametr β je pro dynamické modely poměrně složité odhadnout. Pro odhady v modelech panelových dat se používají různé estimátory. Některé z nich nejsou pro odhady v dynamických modelech vhodné (například OLS estimátor, LSDV estimátor, GLS estimátor).⁶²

⁶⁰ ARLT, Josef a ARLTOVÁ, Markéta. *Ekonomické časové řady*. Praha: Professional Publishing, 2009. s. 11. ISBN 978-80-86946-85-6.

⁶¹ BALTAGI, Badi H. *Econometric Analysis of Panel Data*. 6th ed. Cham: Springer Cham, 2021. s. 187. ISBN 9783030539528.

⁶² BOND, Stephen R. Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal*. 2002, s. 141–162.

Baltagi uvádí následující estimátory od různých autorů, které jsou pro dynamické modely vhodné:

- Arellano-Bond estimátor (1991),
- Arellano-Bover estimátor (1995),
- Blundell-Bond systémový GGM estimátor (1998),
- Keane-Runkle estimátor,
- Limited Information Maximum Likelihood.⁶³

2.4.1 Autoregresivní modely

Autoregresivní modely jsou speciálním případem modelů dynamických. Pokud model mezi své vysvětlující proměnné zahrnuje jednu nebo více zpožděných hodnot závislé proměnné, nazývá se autoregresivní.⁶⁴

Autoregresivní model lze vyjádřit takto:

$$y_{it} = \gamma y_{t-1} + \alpha_i + u_{it} \quad (2.7)$$

kde:

- y_{it} je vysvětlovaná proměnná jednotky i ($i = 1, 2, \dots, n$) a čase t ($t = 1, 2, \dots, T$),
- y_{t-1} vyjadřuje zpožděnou závisle proměnnou,
- α_i vyjadřuje individuální efekty rozdílné pro každé i ,
- u_{it} je reziduální složka pro jednotku i v časovém okamžiku t .

Předpokládá se, že individuální efekty jsou stochastické, tzn. jsou korelovány se zpožděnou závisle proměnnou.⁶⁵

⁶³ BALTAGI, Badi H. *Econometric Analysis of Panel Data*. 6th ed. Cham: Springer Cham, 2021. s. 187. ISBN 9783030539528.

⁶⁴ GUJARATI, Damodar N. a PORTER, Dawn C. *Basic econometrics*. 5th ed. McGraw-Hill/Irwin, 2009. s. 624. ISBN 978-0-07-337577-9.

⁶⁵ BOND, Stephen R. Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal*. 2002, s. 141–162.

3. APLIKACE ANALÝZY PANELOVÝCH DAT

V následující kapitole bude analýza panelových dat aplikována na konkrétní problém z podnikové praxe. Bude zkoumána závislost tržeb na návštěvnosti a počtu zaměstnanců stánků s občerstvením na fotbalovém stadionu. Cílem je zjistit závislosti, které budou pro provozovatele stánku hodnotné.

Nejdříve budou vyhodnoceny modely s jednou vysvětlující proměnnou, a to návštěvností. Následně budou stejné modely vyhodnoceny pro dvě vysvětlující proměnné – návštěvnost a počet zaměstnanců stánku. Modely mezi sebou budou porovnány.

3.1 Charakteristika dat

Datový soubor pro praktickou část této práce obsahuje údaje ze stánků s občerstvením, které se nacházejí na fotbalovém stadionu. Konkrétně jsou uvedené jednotlivé tržby, návštěvnost a počet zaměstnanců deseti stánků, které jsou rozmístěné v rámci celého stadionu. Hodnoty byly zaznamenávány během dvou fotbalových sezón, a to každý zápas 1. české fotbalové ligy, který se na daném stadionu hrál. Vzhledem k tomu, že přesná výše tržeb není pro tuto práci podstatná, hodnoty jsou na žádost poskytovatele dat upraveny pomocí koeficientu, aby data mohla být v této práci zveřejněna. Návštěvnost zápasů je uvedena pro sektory, které spadají pod jednotlivé stánky. Žádné hodnoty v datovém souboru nechybí, jedná se tedy o vyrovnaný panel. Tabulka s výchozími daty je uvedena v příloze 1.

Data obsahují hodnoty závisle proměnné y_{it} , tedy hodnoty tržeb. Hodnoty nezávisle proměnné x_{it1} představují hodnoty návštěvnosti a hodnoty x_{it2} představují počet zaměstnanců. Jedná se tedy o vícerozměrná panelová data.

3.2 Popisné statistiky

V Tabulka 5 jsou shrnuty základní popisné statistiky jednotlivých proměnných napříč všemi objekty. Z tabulky lze vyčíst, že se průměr od mediánu příliš neliší a obě statistiky se nachází přibližně uprostřed mezi minimem a maximem. Platí, že v rámci všech stánků je nejméně 50 % hodnot tržeb nižší nebo rovno 40 880,26 Kč a 50 % hodnot je vyšší než nebo rovno 40 880,26 Kč. Minimální tržba v rámci jednoho stánku byla 8 876 Kč, maximální tržba činila 113 168 Kč. Co se týká návštěvnosti, minimální počet lidí v rámci sektoru byl 206 a maximální počet byl 1 356 lidí.

Tabulka 5 Popisné statistiky proměnných

HODNOTY VŠECH STÁNKŮ				
	průměr	medián	minimum	maximum
tržby	40880,26	40400	8876	113168
návštěvnost	697,65	617,5	206	1356
počet zaměstnanců	9,656	10	4	15

Zdroj: vlastní zpracování

Následující tabulky ukazují popisné statistiky jednotlivých proměnných všech stánků zvlášť.

Tabulka 6 Popisné statistiky jednotlivých stánků – tržby

TRŽBY				
	průměr	medián	minimum	maximum
stánek 1	34434,6	31814	18764	91440
stánek 2	54493,6	51074	40228	86512
stánek 3	50770,4	51238	32628	74556
stánek 4	43302,1	43564	24764	66508
stánek 5	33167,1	24784	8876	113168
stánek 6	48940,3	46928	24960	82520
stánek 7	25505,1	24922	12564	42784
stánek 8	41870,9	40498	20916	60532
stánek 9	43379,8	41852	25836	85532
stánek 10	32938,7	32450	17856	58256

Zdroj: vlastní zpracování

Tabulka 7 Popisné statistiky jednotlivých stánků – návštěvnost

NÁVŠTĚVNOST				
	průměr	medián	minimum	maximum
stánek 1	590,0	594,5	426	731
stánek 2	1107,5	1117	855	1307
stánek 3	972,4	975	690	1078
stánek 4	918,7	860	546	1356
stánek 5	612,7	493,5	299	1322
stánek 6	645,9	543,5	268	1128
stánek 7	405,4	360,5	206	721
stánek 8	621,1	582,5	292	1020
stánek 9	623,3	582,5	310	1035
stánek 10	479,4	454,5	259	736

Zdroj: vlastní zpracování

Tabulka 8 Popisné statistiky jednotlivých stánků – počet zaměstnanců

POČET ZAMĚSTNANCŮ				
	průměr	medián	minimum	maximum
stánek 1	9,7	10	8	10
stánek 2	9,7	10	9	10
stánek 3	9,9	10	8	11
stánek 4	10,1	10	9	12
stánek 5	11,1	11	9	13
stánek 6	13,4	13,5	10	15
stánek 7	4,8	5	4	5
stánek 8	9,4	10	8	10
stánek 9	9,6	10	8	11
stánek 10	8,9	9	8	10

Zdroj: vlastní zpracování

Pokud se průměr výrazně liší od mediánu, znamená to, že se v datech nacházejí extrémní hodnoty. Lze vidět, že v rámci tržeb se průměr od mediánu liší nejvíce u stánku 5, dále u stánků 2, 1 a 6. U návštěvnosti je největší rozdíl mezi průměrem a mediánem u stánků 5 a 6. Jinak se hodnoty příliš neliší. Stejně tak jsou hodnoty průměru a mediánu podobné u počtu zaměstnanců. Díky maximu a minimu je vidět, v jakém rozmezí se tržby, návštěvnost a počet zaměstnanců daných stánků pohybují.

Pro přehlednost jsou průměrné hodnoty všech proměnných shrnuty v Tabulka 9. Jde vidět, že nejnižší průměr tržeb, návštěvnosti i počtu zaměstnanců má stánek 7. Nízké hodnoty u stánku 7 vyplývají z toho, že stánek je menší než ty ostatní.

Nejvyšší hodnoty návštěvnosti a tržeb mají stánky 2 a 3, ale počty zaměstnanců jsou u těchto stánků nižší než například u stánků 5 a 6, které mají nižší návštěvnost. Stánek 5 má také mnohem nižší tržby.

Stánky 2 a 3 mají vysoké hodnoty tržeb a návštěvnosti, protože se nachází na atraktivním místě stadionu, které je vždy fanoušky nejvíce obsazováno.

Tabulka 9 Průměrné hodnoty proměnných

PRŮMĚRNÉ HODNOTY			
	tržby	návštěvnost	počet zaměstnanců
stánek 1	34434,6	590	9,7
stánek 2	54493,6	1107,5	9,7
stánek 3	50770,4	972,4	9,9
stánek 4	43302,1	918,7	10,1
stánek 5	33167,1	612,7	11,1
stánek 6	48940,3	645,9	13,4
stánek 7	25505,1	405,4	4,8
stánek 8	41870,9	621,1	9,4
stánek 9	43379,8	623,3	9,6
stánek 10	32938,7	479,4	8,9

Zdroj: vlastní zpracování

Grafy vývoje jednotlivých proměnných jsou uvedeny v přílohách – příloha 2, příloha 3, příloha 4.

3.3 Aplikace vybraných modelů na jednorozměrná panelová data

V následující podkapitole jsou data zpracována za pomoci jednotlivých modelů, které byly popsány v kapitole 2. Konkrétně je využit spojený regresní model, model s fixními efekty a model s náhodnými efekty. Do modelů jsou zahrnuté pouze proměnné „tržby“ a „návštěvnost“. V kapitole 3.5 jsou následně zpracovány modely i s proměnnou „počet zaměstnanců“. Data jsou zpracována pomocí softwaru Gretl a následně interpretována.

Pro daná data není možné získat dynamické modely, vzhledem k tomu, že matice není pozitivně definitní. Proto v praktické části nejsou dynamické modely uvedeny.

3.3.1 Spojený regresní model

Obrázek 1 ukazuje výstup ze softwaru Gretl po použití metody nejmenších čtverců (Ordinary Least Squares).

Model 1: Hromadné OLS, za použití 360 pozorování
 Zahrnuto 10 průřezových jednotek
 Délka časové řady = 36
 Závisle proměnná: trzby

	koeficient	směr. chyba	t-podíl	p-hodnota	
const	13631.1	1473.18	9.253	2.03e-18	***
navstevnost	39.0585	1.96609	19.87	9.94e-60	***
Střední hodnota závisle proměnné			40880.26		
Sm. odchylka závisle proměnné			14764.95		
Součet čtverců reziduí			3.72e+10		
Sm. chyba regrese			10197.17		
Koeficient determinace			0.524354		
Adjustovaný koeficient determinace			0.523026		
F(1, 358)			394.6609		
P-hodnota(F)			9.94e-60		
Logaritmus věrohodnosti			-3832.567		
Akaikovo kritérium			7669.133		
Schwarzovo kritérium			7676.905		
Hannan-Quinnovo kritérium			7672.224		
rho (koeficient autokorelace)			0.496029		
Durbin-Watsonova statistika			0.969446		

zde je poznámka o zkratkách statistik modelu

Obrázek 1 Spojený regresní model 1

Zdroj: vlastní zpracování (software Gretl)

Spojený regresní model lze podle výstupu ze softwaru Gretl zapsat takto:

$$y_{it} = 13631,1 + 39,0585x_{it1} + u_{it} \quad (3.1)$$

U spojeného regresního modelu určuje kvalitu modelu koeficient determinace a adjustovaný koeficient determinace, který počítá s tím, že do modelu nejsou zahrnuty případné nadbytečné proměnné. Hodnoty koeficientů se pohybují v rozmezí 0–1. Čím vyšší koeficienty jsou, tím lépe model predikuje hodnoty závisle proměnné. A je tedy více užitečný. V tomto případě jsou hodnoty obou koeficientů podobné. Hodnota koeficientu determinace je 0,524354 a hodnota adjustovaného koeficientu determinace je 0,523026. Hodnoty se tedy nachází přibližně v půlce intervalu.

Z modelu lze dle kladného znaménka u koeficientu proměnné „návštěvnost“ určit, že pokud stoupá návštěvnost, stoupají i tržby. Pokud chce firma zvýšit tržby, měla by usilovat o vyšší návštěvnost, což je zřejmé. Samozřejmě při tom musí brát v úvahu kapacitu jednotlivých sektorů. V tomto modelu je nezávisle proměnná významná. To lze určit pomocí p-hodnoty a přidělených hvězdiček.

P-hodnota přísluší testu významnosti daného koeficientu, tedy hypotézy $H_0: \beta = 0$. Výsledné p-hodnoty $\leq 0,05$ vedou k zamítnutí hypotézy na hladině významnosti $\alpha = 0,05$. To znamená, že koeficient je v modelu významný.

3.3.2 Model s fixními efekty

Obrázek 2 ukazuje výstup ze softwaru Gretl po použití modelu s fixními efekty bez zahrnutí časových individuálních efektů.

```
Model 2: Pevné efekty, za použití 360 pozorování
Zahrnuto 10 průřezových jednotek
Délka časové řady = 36
Závisle proměnná: trzby

      koeficient  směr. chyba  t-podíl  p-hodnota
-----
const      8394.33      2004.09      4.189    3.56e-05 ***
navstevnost  46.5648      2.79096     16.68    2.30e-46 ***

Střední hodnota závisle proměnné  40880.26
Sm. odchylka závisle proměnné     14764.95
Součet čtverců reziduí              2.83e+10
Sm. chyba regrese                    9002.295
LSDV R-squared                       0.638612
Within R-squared                     0.443701
LSDV F(10, 349)                      61.67218
P-hodnota(F)                         5.06e-71
Logaritmus věrohodnosti              -3783.117
Akaikovo kritérium                   7588.233
Schwarzovo kritérium                 7630.980
Hannan-Quinnovo kritérium            7605.230
rho (koeficient autokorelace)        0.323345
Durbin-Watsonova statistika          1.298399
zde je poznámka o zkratkách statistik modelu

Joint test on named regressors -
Testovací statistika: F(1, 349) = 278.361
s p-hodnotou = P(F(1, 349) > 278.361) = 2.2982e-46

Test pro různé intercepty mezi skupinami -
Nulová hypotéza: Skupiny mají společný intercept
Testovací statistika: F(9, 349) = 12.2602
s p-hodnotou = P(F(9, 349) > 12.2602) = 6.92321e-17
```

Obrázek 2 Model s fixními efekty 1

Zdroj: vlastní zpracování (software Gretl)

Kvalitu modelu, v případě použití modelů s individuálními efekty, posuzují informační kritéria. Mezi informační kritéria patří Akaikovo kritérium, Schwarzovo kritérium a Hannan-Quinnovo kritérium. Čím nižší je hodnota informačního kritéria, tím je kvalita modelu vyšší. Jejich hodnoty u modelu s fixními efekty jsou:

- Akaikovo kritérium – 7 588,233,
- Schwarzovo kritérium – 7 630,980,
- Hannan-Quinnovo kritérium – 7 605,230.

Díky p-hodnotám lze vyčíst, že v tomto modelu má proměnná „návštěvnost“ na model výrazný vliv.

Model s fixními efekty bez zahrnutí časových individuálních efektů lze podle výsledků ze softwaru Gretl zapsat takto:

$$y_{it} = 8394,33 + 46,5648x_{it1} + u_{it} \quad (3.2)$$

Pomocí testu pro různé intercepty mezi skupinami lze určit, zda je model s fixními efekty pro model vhodnější než spojený regresní model. Nulová hypotéza o společném interceptu pro všechny průřezové jednotky se zamítá, protože je p-hodnota příliš nízká. To značí, že je vhodnější použití modelu s fixními efekty než spojeného regresního modelu.

Obrázek 3 ukazuje výstup ze softwaru Gretl po použití modelu s fixními efekty se zahrnutím časových individuálních efektů.

Model 3: Pevné efekty, za použití 360 pozorování
 Zahrnuto 10 průřezových jednotek
 Délka časové řady = 36
 Závisle proměnná: trzby

	koeficient	směr. chyba	t-podíl	p-hodnota	
const	24587.6	3197.42	7.690	1.91e-13	***
navstevnost	25.4574	3.61535	7.041	1.20e-11	***
dt_2	3664.74	3304.61	1.109	0.2683	
dt_3	-571.586	3536.23	-0.1616	0.8717	
dt_4	-2622.42	3286.74	-0.7979	0.4255	
dt_5	-479.448	3408.91	-0.1406	0.8882	
dt_6	2072.25	3512.74	0.5899	0.5557	
dt_7	-2840.66	3292.21	-0.8628	0.3889	
dt_8	-12183.1	3287.63	-3.706	0.0002	***
dt_9	-10936.7	3286.76	-3.328	0.0010	***
dt_10	-8103.65	3288.30	-2.464	0.0143	**
dt_11	-11353.7	3287.40	-3.454	0.0006	***
dt_12	-4907.62	3371.24	-1.456	0.1465	
dt_13	-6420.35	3289.01	-1.952	0.0518	*
dt_14	-7408.95	3289.87	-2.252	0.0250	**
dt_15	-4401.44	3293.60	-1.336	0.1824	
dt_16	-5400.26	3286.72	-1.643	0.1014	
dt_17	-9856.20	3290.47	-2.995	0.0030	***
dt_18	690.198	3290.47	0.2098	0.8340	
dt_19	-525.246	3290.08	-0.1596	0.8733	
dt_20	81.5052	3431.29	0.02375	0.9811	
dt_21	1820.85	3329.92	0.5468	0.5849	
dt_22	-345.922	3323.77	-0.1041	0.9172	
dt_23	-1485.41	3305.48	-0.4494	0.6535	
dt_24	-2081.89	3291.38	-0.6325	0.5275	
dt_25	-4840.61	3291.48	-1.471	0.1424	
dt_26	-3653.77	3288.95	-1.111	0.2675	
dt_27	-11918.2	3355.38	-3.552	0.0004	***
dt_28	5552.89	3304.57	1.680	0.0939	*
dt_29	6910.34	3304.61	2.091	0.0373	**
dt_30	14760.6	3597.67	4.103	5.21e-05	***
dt_31	3643.66	3334.99	1.093	0.2754	
dt_32	19611.3	3581.39	5.476	8.93e-08	***
dt_33	3103.79	3305.91	0.9389	0.3485	
dt_34	-771.696	3290.90	-0.2345	0.8148	
dt_35	-3668.24	3292.48	-1.114	0.2661	
dt_36	2028.54	3288.26	0.6169	0.5377	

Střední hodnota závisle proměnné 40880.26
 Sm. odchylka závisle proměnné 14764.95
 Součet čtverců reziduí 1.70e+10
 Sm. chyba regrese 7349.306
 LSDV R-squared 0.783298
 Within R-squared 0.666421
 LSDV F(45, 314) 25.22204
 P-hodnota(F) 8.74e-80
 Logaritmus věrohodnosti -3691.060
 Akaikovo kritérium 7474.120
 Schwarzovo kritérium 7652.881
 Hannan-Quinnovo kritérium 7545.198
 rho (koeficient autokorelace) 0.220640
 Durbin-Watsonova statistika 1.473829
 zde je poznámka o zkratkách statistik modelu

Joint test on named regressors -
 Testovací statistika: $F(1, 314) = 49.5824$
 s p-hodnotou = $P(F(1, 314) > 49.5824) = 1.20439e-11$

Test pro různé intercepty mezi skupinami -
 Nulová hypotéza: Skupiny mají společný intercept
 Testovací statistika: $F(9, 314) = 16.9635$
 s p-hodnotou = $P(F(9, 314) > 16.9635) = 8.83956e-23$

Wald joint test on time dummies -
 Nulová hypotéza: No time effects
 Asymptotická testovací statistika: $\text{Chí-kvadrát}(35) = 209.648$
 s p-hodnotou = 8.99485e-27

Obrázek 3 Model s fixními efekty se zahrnutím individuálních časových efektů 1

Zdroj: vlastní zpracování (software Gretl)

Hodnoty informačních kritérií, které posuzují kvalitu modelu jsou u tohoto modelu:

- Akaikovo kritérium – 7 474,120,
- Schwarzovo kritérium – 7 652,881,
- Hannan-Quinnovo kritérium – 7 545,198.

Hodnoty Akaikova kritéria a Hannan-Quinnova kritéria jsou při zahrnutí časových efektů nižší. Hodnota Schwarzova kritéria je jen nepatrně vyšší.

Proměnná návštěvnost je v modelu s fixními efekty včetně individuálních časových efektů velmi významná. Hodnoty časových efektů nemají klesající ani stoupající trend. Některé z časových efektů mají na model významný vliv, a to časový efekt 8, 9, 11, 17, 27, 30 a 32.

Za pomoci p-hodnoty testu „Wald joint test on time dummies“, můžeme zamítnout hypotézu o nepotřebnosti časových individuálních efektů, protože p-hodnota je příliš nízká. To znamená, že pro daná data je vhodnější zahrnutí individuálních časových efektů do modelu s fixními efekty.

3.3.3 Model s náhodnými efekty

Stejně jako u modelu s fixními efekty, se kvalita modelu s náhodnými efekty určuje podle informačních kritérií. Pro tento model jsou hodnoty kritérií následující:

- Akaikovo kritérium – 7 678,717,
- Schwarzovo kritérium – 7 686,490,
- Hannan-Quinnovo kritérium – 7 681,808.

Dle těchto hodnot lze určit, že pro daná data je vhodnější model fixních efektů se zahrnutím časových individuálních efektů než model s náhodnými efekty, protože hodnoty všech informačních kritérií jsou u tohoto modelu vyšší.

Obrázek 4 ukazuje výstup ze softwaru Gretl po použití modelu s náhodnými efekty, podle něj lze model zapsat takto:

$$y_{it} = 24587,6 + 25,4574x_{it1} + (\alpha + \varepsilon_i) + u_{it} \quad (3.3)$$

Model 4: Náhodné efekty (GLS), za použití 360 pozorování
 Zahrnuto 10 průřezových jednotek
 Délka časové řady = 36
 Závisle proměnná: trzby

	koeficient	směr. chyba	z	p-hodnota	
const	9368.15	2480.40	3.777	0.0002	***
navstevnost	45.1689	2.63138	17.17	4.82e-66	***
Střední hodnota závisle proměnné		40880.26			
Sm. odchylka závisle proměnné		14764.95			
Součet čtverců reziduí		3.82e+10			
Sm. chyba regrese		10319.41			
Logaritmus věrohodnosti		-3837.359			
Akaikovo kritérium		7678.717			
Schwarzovo kritérium		7686.490			
Hannan-Quinnovo kritérium		7681.808			
rho (koeficient autokorelace)		0.323345			
Durbin-Watsonova statistika		1.298399			

zde je poznámka o zkratkách statistik modelu

Rozptyl 'mezi' = 2.54791e+07
 Rozptyl 'uvnitř' = 8.10413e+07
 theta použitá pro částečné odstranění střední hodnoty = 0.715079
 corr(y,yhat)^2 = 0.524354

Joint test on named regressors -
 Asymptotická testovací statistika: Chí-kvadrát(1) = 294.653
 s p-hodnotou = 4.8159e-66

Breusch-Paganův test -
 Nulová hypotéza: Rozptyl chyb příslušejících jednotkám = 0
 Asymptotická testovací statistika: Chí-kvadrát(1) = 257.859
 s p-hodnotou = 5.02645e-58

Hausmanův test -
 Nulová hypotéza: GLS odhady jsou konzistentní
 Asymptotická testovací statistika: Chí-kvadrát(1) = 2.2119
 s p-hodnotou = 0.13695

Obrázek 4 Model s náhodnými efekty 1

Zdroj: vlastní zpracování (software Gretl)

Obrázek 5 ukazuje výstup ze softwaru Gretl po použití modelu s náhodnými efekty a se zahrnutím časových individuálních efektů.

Lze z něj vyčíst, že proměnná „navstevnost“ je tomto modelu velmi významná. Hodnoty časových efektů nemají klesající ani stoupající trend. Časové efekty, které jsou velmi významné – tedy mají tři hvězdičky, jsou stejné jako u modelu s fixními efekty. Důvodem mohou být opět hodnoty tržeb zapříčiněné atraktivitou zápasu.

Podle Hausmanova testu lze zamítnout jeho nulovou hypotézu, že odhady jsou konzistentní. To znamená, že model s náhodnými efekty není lepší než model s fixními efekty.

Model 5: Náhodné efekty (GLS), za použití 360 pozorování
 Zahrnuto 10 průřezových jednotek
 Délka časové řady = 36
 Závisle proměnná: trzby

	koeficient	směr. chyba	z	p-hodnota	
const	23630.9	3463.02	6.824	8.87e-12	***
navstevnost	27.0324	3.27893	8.244	1.66e-16	***
dt_2	3515.12	3301.78	1.065	0.2871	
dt_3	-1139.99	3493.58	-0.3263	0.7442	
dt_4	-2615.97	3287.09	-0.7958	0.4261	
dt_5	-873.502	3387.89	-0.2578	0.7965	
dt_6	1532.20	3474.03	0.4410	0.6592	
dt_7	-2923.50	3291.58	-0.8882	0.3744	
dt_8	-12149.2	3287.81	-3.695	0.0002	***
dt_9	-10944.5	3287.10	-3.330	0.0009	***
dt_10	-8148.22	3288.37	-2.478	0.0132	**
dt_11	-11324.4	3287.62	-3.445	0.0006	***
dt_12	-5234.42	3356.73	-1.559	0.1189	
dt_13	-6473.90	3288.95	-1.968	0.0490	**
dt_14	-7471.79	3289.66	-2.271	0.0231	**
dt_15	-4494.21	3292.73	-1.365	0.1723	
dt_16	-5403.57	3287.06	-1.644	0.1002	
dt_17	-9787.69	3290.15	-2.975	0.0029	***
dt_18	758.709	3290.15	0.2306	0.8176	
dt_19	-590.134	3289.83	-0.1794	0.8576	
dt_20	-347.828	3406.42	-0.1021	0.9187	
dt_21	1587.91	3322.64	0.4779	0.6327	
dt_22	-561.533	3317.57	-0.1693	0.8656	
dt_23	-1638.65	3302.50	-0.4962	0.6198	
dt_24	-2158.27	3290.90	-0.6558	0.5119	
dt_25	-4917.79	3290.98	-1.494	0.1351	
dt_26	-3706.69	3288.90	-1.127	0.2597	
dt_27	-11624.0	3343.64	-3.476	0.0005	***
dt_28	5403.43	3301.75	1.637	0.1017	
dt_29	6760.72	3301.78	2.048	0.0406	**
dt_30	14123.2	3544.80	3.984	6.77e-05	***
dt_31	3397.34	3326.82	1.021	0.3072	
dt_32	18991.6	3531.22	5.378	7.52e-08	***
dt_33	2948.81	3302.85	0.8928	0.3720	
dt_34	-843.986	3290.50	-0.2565	0.7976	
dt_35	-3583.35	3291.81	-1.089	0.2763	
dt_36	1984.60	3288.33	0.6035	0.5462	

Střední hodnota závisle proměnné 40880.26
 Sm. odchylka závisle proměnné 14764.95
 Součet čtverců reziduí 2.58e+10
 Sm. chyba regrese 8929.345
 Logaritmus věrohodnosti -3766.808
 Akaikovo kritérium 7607.617
 Schwarzovo kritérium 7751.403
 Hannan-Quinnovo kritérium 7664.789
 rho (koeficient autokorelace) 0.220640
 Durbin-Watsonova statistika 1.473829

zde je poznámka o zkratkách statistik modelu

Rozptyl 'mezi' = 2.62299e+07
 Rozptyl 'uvnitř' = 5.40123e+07
 theta použitá pro částečné odstranění střední hodnoty = 0.767395
 corr(y,yhat)^2 = 0.672527

Joint test on named regressors -
 Asymptotická testovací statistika: Chí-kvadrát(1) = 67.9678
 s p-hodnotou = 1.66187e-16

Wald joint test on time dummies -
 Nulová hypotéza: No time effects
 Asymptotická testovací statistika: Chí-kvadrát(35) = 210.815
 s p-hodnotou = 5.49417e-27

Breusch-Paganův test -
 Nulová hypotéza: Rozptyl chyb příslušejících jednotkám = 0
 Asymptotická testovací statistika: Chí-kvadrát(1) = 564.911
 s p-hodnotou = 7.18329e-125

Hausmanův test -
 Nulová hypotéza: GLS odhady jsou konzistentní
 Asymptotická testovací statistika: Chí-kvadrát(1) = 1.1945
 s p-hodnotou = 0.274423

Obrázek 5 Model s náhodnými efekty se zahrnutím individuálních časových efektů 1

Zdroj: vlastní zpracování (software Gretl)

3.4 Výběr optimálního modelu

U spojeného regresního modelu určují kvalitu modelu koeficient determinace a adjustovaný koeficient determinace. U modelů, které obsahují individuální efekty se určuje kvalita modelu podle informačních kritérií. Vzhledem k této skutečnosti je třeba určit kritérium, které slouží pro porovnání všech modelů mezi sebou, a to je součet čtverců reziduí. Čím menší hodnotu má, tím více model vystihuje data. V Tabulka 10 je uveden přehled součtu čtverců reziduí pro všechny výše zpracované modely.

Tabulka 10 Součty čtverců reziduí pro jednotlivé modely

model	součet čtverců reziduí
1 – Spojený regresní model	37 200 000 000
2 – Model s fixními efekty	28 300 000 000
3 – Model s fixními efekty se zahrnutím individuálních časových efektů	17 000 000 000
4 – Model s náhodnými efekty	38 200 000 000
5 – Model s náhodnými efekty se zahrnutím individuálních časových efektů	25 800 000 000

Zdroj: vlastní zpracování

Z tabulky vyplývá, že nejvhodnější model je model s fixními efekty se zahrnutím individuálních časových efektů, protože má nejnižší hodnotu součtu čtverců reziduí. Hausmanův test také potvrdil, že je pro daná data vhodnější model fixních efektů než model s náhodnými efekty. Zároveň byla zamítnuta hypotéza o nepotřebnosti časových individuálních efektů, a tedy je tento model vhodnější než model fixních efektů bez zahrnutí časových individuálních proměnných.

3.4.1 Analýza výsledků vybraného modelu

V následující podkapitole bude rozebrán model s fixními efekty se zahrnutím individuálních časových efektů, který vyšel jako nejvhodnější pro daná data.

Model lze zapsat takto:

$$y_{it} = 24587,6 + 25,4574x_{it1} + \varepsilon_i + \lambda_t + u_{it} \quad (3.4)$$

kde:

ε_i vyjadřuje individuální efekty stánků,

λ_t vyjadřuje individuální časové efekty.

Z rovnice lze vyčíst, že pokud se návštěvnost zvýší o jednoho člověka, potom tržba stánku vzroste v průměru o 25,4574 Kč. Pokud by vliv všech nezávislých proměnných v modelu byl nulový, bude tržba stánku rovna 24 587,6 Kč.

Jak již bylo zmíněno, kladné znaménko u koeficientu proměnné „návštěvnost“ vyjadřuje, že pokud stoupá návštěvnost, stoupají i tržby.

Parametr u_{it} zahrnuje nepozorované efekty, které na data mohou působit. Mezi nepozorované efekty v tomto případě může patřit například:

- atraktivita soupeře,
- čas zápasu,
- počasí v den konání zápasu,
- nabídka kulturních a sportovních akcí v daném městě v době konání zápasu,
- umístění týmu ve fotbalové soutěži.

Tabulka 11 shrnuje individuální časové efekty a jejich významnost.

Tabulka 11 Koeficienty individuálních časových efektů

zápas	koeficient časového efektu	významnost	zápas	koeficient časového efektu	významnost
1	–		19	–525,246	
2	3664,74		20	81,5052	
3	–571,586		21	1820,85	
4	–2622,42		22	–345,922	
5	–479,448		23	–1485,41	
6	2072,25		24	–2081,89	
7	–2840,66		25	–4840,61	
8	–12183,1	***	26	–3653,77	
9	–10936,7	***	27	–11918,2	***
10	–8103,65	**	28	5552,89	*
11	–11353,7	***	29	6910,34	**
12	–4907,62		30	14760,6	***
13	–6420,35	*	31	3643,66	
14	–7408,95	**	32	19611,3	***

zápas	koeficient časového efektu	významnost	zápas	koeficient časového efektu	významnost
15	-4401,44		33	3103,79	
16	-5400,26		34	-771,696	
17	-9856,20	***	35	-3668,24	
18	690,198		36	2028,54	

Zdroj: vlastní zpracování

Velmi významné jsou časové efekty 8, 9, 11, 17, 27, 30 a 32. To může být zapříčiněno výší tržeb. V tabulce níže jsou uvedeny celkové tržby všech stánků za jednotlivé zápasy. Jak můžeme vidět v Tabulka 12, v zápasech číslo 8, 9, 11, 17 a 27 byly celkové tržby nejnižší, naopak v zápasech 30 a 32 byly tržby nejvyšší. Rozdíl mezi nejnižší a nejvyšší tržbou je značný.

Vysoké tržby u posledních dvou zmiňovaných zápasů mohou být způsobeny atraktivitou soupeře a časem, ve kterých se zápasy hrály – 18:00. Oba zápasy se také hrály v neděli. Naopak nízké tržby mohou být způsobeny neatraktivitou soupeře nebo časem začátku utkání. Všechny zápasy se hrály od 15:00 hodin.

Tabulka 12 Přehled celkových tržeb jednotlivých zápasů

zápas	celkové tržby	zápas	celkové tržby
1	400 504 Kč	19	405 740 Kč
2	461 336 Kč	20	470 716 Kč
3	486 664 Kč	21	456 364 Kč
4	373 236 Kč	22	431 896 Kč
5	459 404 Kč	23	410 420 Kč
6	508 520 Kč	24	392 032 Kč
7	385 488 Kč	25	364 572 Kč
8	273 200 Kč	26	372 520 Kč
9	292 384 Kč	27	233 768 Kč
10	326 672 Kč	28	480 192 Kč
11	282 232 Kč	29	493 792 Kč
12	404 252 Kč	30	651 136 Kč
13	344 956 Kč	31	476 756 Kč
14	336 572 Kč	32	696 792 Kč
15	371 484 Kč	33	456 592 Kč
16	347 036 Kč	34	404 472 Kč
17	290 868 Kč	35	350 100 Kč
18	396 332 Kč	36	427 892 Kč

Zdroj: vlastní zpracování

V Tabulka 13 jsou seskupeny zápasy, které byly odehrány se stejným soupeřem. K nim jsou uvedené koeficienty individuálních časových efektů.

Tabulka 13 Hodnoty individuálních časových efektů v rámci zápasů se stejným soupeřem 1

skupiny	hodnoty koeficientů časových efektů			
2, 16, 26, 35	3664,74	-5400,26	-3653,77	-3668,24
12, 18, 22	2072,25	14760,6		
6, 30	-4907,62	690,198	-345,922	
5, 20, 36	-479,448	81,5052	2028,54	
8, 19	-12183,1	-525,246		
10, 31	-8103,65	3643,66		
14, 29	-7408,95	6910,34		
13, 24	-6420,35	-2081,89		
4, 28, 34	-2622,42	5552,89	-771,696	
15, 33	-4401,44	3103,79		
9, 27	-10936,7	-11918,2		
11, 17, 21	-11353,7	-9856,2	1820,85	
3, 32	-571,586	19611,3		

Zdroj: vlastní zpracování

Hodnoty časových koeficientů v rámci skupin jsou velmi rozdílné. Nelze předpokládat stejný časový efekt u zápasů se stejným soupeřem.

V Tabulka 14 jsou seskupeny zápasy podle stejného času začátku utkání. Podle hodnot koeficientů časových efektů, které k nim připadají, nelze říct, že by byl individuální časový efekt podobný.

Tabulka 14 Hodnoty individuálních časových efektů v rámci zápasů se stejným časem začátku utkání 1

	čas utkání			
	15:00	17:00	18:00	19:00
hodnoty koeficientů časových efektů	-12183,1	3664,74	2072,25	-571,586
	-10936,7	-2840,66	-8103,65	-2622,42
	-11353,7	2028,54	-4907,62	-479,448
	-4401,44		-5400,26	-6420,35
	-9856,2		-525,246	-7408,95
	81,5052		-2081,89	690,198
	1820,85		5552,89	
	-345,922		6910,34	
	-1485,41		14760,6	
	-4840,61		19611,3	
	-3653,77			
	-11918,2			
	3643,66			
	3103,79			
	-771,696			
	-3668,24			

Zdroj: vlastní zpracování

Hodnoty časových koeficientů jsou dále seskupeny podle dne v týdnu, kdy se zápas hrál (viz Tabulka 15). Ani z této tabulky nelze usuzovat, že by měl konkrétní den konání zápasu podobný časový efekt.

Tabulka 15 Hodnoty individuálních časových efektů v rámci zápasů se stejným dnem konání 1

	den konání zápasu		
	středa	sobota	neděle
	3664,74	-2622,42	-571,586
	-2840,66	-479,448	-10936,7
		2072,25	-4907,62
		-12183,1	-7408,95
		-8103,65	-4401,44
		-11353,7	-525,246
		-6420,35	81,5052
		-5400,26	1820,85
		-9856,2	-345,922
		5552,89	-1485,41
		6910,34	-2081,89
		3643,66	-4840,61
		-3668,24	-3653,77
		2028,54	-11918,2
			14760,6
			19611,3
			3103,79
			-771,696

Zdroj: vlastní zpracování

V Tabulka 16 jsou uvedeny konstanty vztažené k jednotkám, jejich průměr je potom konstantou závislé proměnné v celkovém modelu.

Tabulka 16 Konstanty vztažené k jednotkám 1

číslo stánku	konstanta
stánek 1	20 882,35
stánek 2	27 766,42
stánek 3	27 484,31
stánek 4	21 381,48
stánek 5	19 036,46
stánek 6	33 966,05
stánek 7	16 651,22
stánek 8	27 525,97
stánek 9	28 978,99
stánek 10	22 202,34

Zdroj: vlastní zpracování

Z Tabulka 16 lze vyčíst tržby stánků, kdyby na ně návštěvnost neměla žádný vliv. Vidíme, že nejvyšší tržbu by měl stánek 6, následovaly by stánky 9, 2, 8 a 3. Nejnižší tržbu by měl stánek 7.

Při porovnání konstant s průměrnou tržbou se shodují nejvyšší hodnoty u stánku 7. Stánek má nejvyšší tržby a měl by je i v případě, že by návštěvnost neměla žádný vliv. Naopak nejvyšší hodnoty průměru tržeb a konstant mají různé stánky. Největší rozdíl je u stánků číslo 2 a 3. Znamená to, že na tyto stánky má návštěvnost velký vliv. Velký vliv má návštěvnost také na stánek 4. Vyšší návštěvnost zvyšuje jejich tržby.

Tabulka 17 Porovnání průměrných tržeb a konstant 1

	průměrné tržby	konstanty vztažené k jednotkám
stánek 1	34434,6	20882,35
stánek 2	54493,6	27766,42
stánek 3	50770,4	27484,31
stánek 4	43302,1	21381,48
stánek 5	33167,1	19036,46
stánek 6	48940,3	33966,05
stánek 7	25505,1	16651,22
stánek 8	41870,9	27525,97
stánek 9	43379,8	28978,99
stánek 10	32938,7	22202,34

Zdroj: vlastní zpracování

3.5 Aplikace vybraných modelů na vícerozměrná panelová data

V následující podkapitole jsou aplikovány stejné modely jako v podkapitole 3.3. Vysvětlující proměnné zahrnují kromě návštěvnosti ještě počet zaměstnanců daných stánků.

Ani po přidání proměnné „zaměstnanci“ není možné získat dynamické modely, vzhledem k tomu, že matice není pozitivně definitní. Ani v této části dynamické modely nejsou uvedeny.

3.5.1 Spojený regresní model

Obrázek 6 ukazuje výstup ze softwaru Gretl po použití metody nejmenších čtverců.

Model 6: Hromadné OLS, za použití 360 pozorování
 Zahrnuto 10 průřezových jednotek
 Délka časové řady = 36
 Závisle proměnná: trzby

	koeficient	směr. chyba	t-podíl	p-hodnota	
const	4014.74	2481.25	1.618	0.1065	
navstevnost	36.7862	1.96874	18.69	7.98e-55	***
zamestnanci	1160.12	244.557	4.744	3.04e-06	***
Střední hodnota závisle proměnné		40880.26			
Sm. odchylka závisle proměnné		14764.95			
Součet čtverců reziduí		3.50e+10			
Sm. chyba regrese		9904.061			
Koeficient determinace		0.552558			
Adjustovaný koeficient determinace		0.550052			
F(2, 357)		220.4348			
P-hodnota(F)		4.53e-63			
Logaritmus věrohodnosti		-3821.564			
Akaikovo kritérium		7649.127			
Schwarzovo kritérium		7660.786			
Hannan-Quinnovo kritérium		7653.763			
rho (koeficient autokorelace)		0.463017			
Durbin-Watsonova statistika		1.029867			

zde je poznámka o zkratkách statistik modelu

Obrázek 6 Spojený regresní model 2

Zdroj: vlastní zpracování (software Gretl)

Spojený regresní model lze podle výstupu ze softwaru Gretl zapsat takto:

$$y_{it} = 4014,74 + 36,7862x_{it1} + 1160,12x_{it2} + u_{it} \quad (3.5)$$

Hodnota koeficientu determinace je 0,552558 a hodnota adjustovaného koeficientu determinace je 0,550052. V porovnání se spojeným regresním modelem bez zahrnutí proměnné „zaměstnanci“ se hodnoty koeficientů příliš neliší. Jsou jen lehce vyšší. Kvalita spojeného regresního modelu se zahrnutím druhé vysvětlující proměnné je tedy nepatrně vyšší.

Za pomoci p-hodnoty lze určit, že v rámci spojeného regresního modelu jsou obě proměnné velmi významné. Obě vysvětlující proměnné mají koeficient s kladným znaménkem. Znamená to, že pokud stoupne návštěvnost o jednoho člověka, tržba stánku stoupne v průměru o 36,7862 Kč. Pokud se obsluha stánku rozšíří o jednoho zaměstnance, tržba stánku stoupne o 1 160,12 Kč.

3.5.2 Model s fixními efekty

Obrázek 7 ukazuje výstup ze softwaru Gretl po použití modelu s fixními efekty bez zahrnutí časových individuálních efektů.

Model 7: Pevné efekty, za použití 360 pozorování
 Zahrnuto 10 průřezových jednotek
 Délka časové řady = 36
 Závisle proměnná: trzby

	koeficient	směr. chyba	t-podíl	p-hodnota	
const	-892.849	5515.39	-0.1619	0.8715	
navstevnost	46.6251	2.78215	16.76	1.24e-46	***
zamestnanci	957.493	530.020	1.807	0.0717	*
Střední hodnota závisle proměnné		40880.26			
Sm. odchylka závisle proměnné		14764.95			
Součet čtverců reziduí		2.80e+10			
Sm. chyba regrese		8973.243			
LSDV R-squared		0.641970			
Within R-squared		0.448870			
LSDV F(11, 348)		56.72593			
P-hodnota(F)		8.25e-71			
Logaritmus věrohodnosti		-3781.436			
Akaikovo kritérium		7586.873			
Schwarzovo kritérium		7633.506			
Hannan-Quinnovo kritérium		7605.415			
rho (koeficient autokorelace)		0.315936			
Durbin-Watsonova statistika		1.308274			

zde je poznámka o zkratkách statistik modelu

Joint test on named regressors -
 Testovací statistika: F(2, 348) = 141.715
 s p-hodnotou = P(F(2, 348) > 141.715) = 9.51128e-46

Test pro různé intercepty mezi skupinami -
 Nulová hypotéza: Skupiny mají společný intercept
 Testovací statistika: F(9, 348) = 9.65628
 s p-hodnotou = P(F(9, 348) > 9.65628) = 3.44606e-13

Obrázek 7 Model s fixními efekty 2

Zdroj: vlastní zpracování (software Gretl)

Výše hodnot informačních kritérií, které určují kvalitu modelu je:

- Akaikovo kritérium – 7 586,873,
- Schwarzovo kritérium –7 633,506,
- Hannan-Quinnovo kritérium –7 605,415.

Hodnoty informačních kritérií jsou velmi podobné jako hodnoty u modelu s fixními efekty bez zahrnutí vysvětlující proměnné „zaměstnanci“. Modely jsou podobně kvalitní.

Proměnná „návštěvnost“ je vyhodnocena jako velmi významná. Výše p-hodnoty u proměnné „zaměstnanci“ je v tomto modelu 0,0717. Protože je p-hodnota > 0,05, je proměnná „zaměstnanci“ vyhodnocena jako málo významná.

Model s fixními efekty bez zahrnutí časových individuálních efektů lze podle výsledků ze softwaru Gretl zapsat takto:

$$y_{it} = -892,849 + 46,6251x_{it1} + 957,493x_{it2} + u_{it} \quad (3.6)$$

Nulová hypotéza o společném interceptu pro všechny průřezové jednotky se zamítá, protože je p-hodnota příliš nízká. Je tedy vhodnější použít model fixních efektů než spojený regresní model.

Obrázek 8 ukazuje výstup ze softwaru Gretl po použití modelu s fixními efekty se zahrnutím časových individuálních efektů.

Z obrázku lze vyčíst, že hodnoty informačních kritérií jsou následující:

- Akaikovo kritérium – 7 472,318,
- Schwarzovo kritérium – 7 654,965,
- Hannan-Quinnovo kritérium – 7 544,942.

Hodnoty jsou velmi podobné jako u modelu s fixními efekty bez zahrnutí proměnné „zaměstnanci“. Nelze tedy říct, že by byl model kvalitnější. Významnost vysvětlující proměnné „návštěvnost“ je vysoká, na rozdíl od proměnné „zaměstnanci“, která má, stejně jako v předchozím modelu, nízkou významnost.

Hodnoty Akaikova kritéria a Hannan-Quinnova kritéria jsou při zahrnutí časových efektů nižší. Hodnota Schwarzova kritéria je jen nepatrně vyšší.

Hypotézu o nepotřebnosti časových individuálních efektů lze na základě p-hodnoty zamítnout. Tedy pro daná data je vhodnější zahrnutí individuálních časových efektů do modelu s fixními efekty.

Model 8: Pevné efekty, za použití 360 pozorování
 Zahrnuto 10 průřezových jednotek
 Délka časové řady = 36
 Závisle proměnná: trzby

	koeficient	směr. chyba	t-podíl	p-hodnota	
const	15298.1	6010.04	2.545	0.0114	**
navstevnost	24.9235	3.61395	6.896	2.95e-11	***
zamestnanci	1131.03	620.494	1.823	0.0693	*
dt_2	3715.46	3292.57	1.128	0.2600	
dt_3	-378.903	3524.81	-0.1075	0.9145	
dt_4	-2511.51	3275.22	-0.7668	0.4438	
dt_5	-345.868	3397.17	-0.1018	0.9190	
dt_6	2255.32	3501.26	0.6441	0.5200	
dt_7	-2812.58	3280.13	-0.8575	0.3918	
dt_8	-13325.6	3334.97	-3.996	8.05e-05	***
dt_9	-11952.1	3321.70	-3.598	0.0004	***
dt_10	-9785.08	3403.59	-2.875	0.0043	***
dt_11	-13286.4	3442.65	-3.859	0.0001	***
dt_12	-6945.79	3540.07	-1.962	0.0506	*
dt_13	-8211.85	3421.13	-2.400	0.0170	**
dt_14	-8971.09	3387.96	-2.648	0.0085	***
dt_15	-5840.34	3375.10	-1.730	0.0845	*
dt_16	-6982.58	3387.74	-2.061	0.0401	**
dt_17	-10558.0	3300.90	-3.199	0.0015	***
dt_18	-916.471	3394.79	-0.2700	0.7874	
dt_19	-1860.49	3358.83	-0.5539	0.5800	
dt_20	-1130.19	3482.70	-0.3245	0.7458	
dt_21	-22.9448	3468.45	-0.006615	0.9947	
dt_22	-1969.38	3429.22	-0.5743	0.5662	
dt_23	-3356.21	3449.54	-0.9729	0.3313	
dt_24	-3978.75	3440.43	-1.156	0.2484	
dt_25	-6058.59	3346.76	-1.810	0.0712	*
dt_26	-5445.48	3421.11	-1.592	0.1125	
dt_27	-13827.5	3503.31	-3.947	9.78e-05	***
dt_28	3793.91	3430.92	1.106	0.2697	
dt_29	5264.52	3414.01	1.542	0.1241	
dt_30	13053.9	3704.71	3.524	0.0005	***
dt_31	1804.41	3472.56	0.5196	0.6037	
dt_32	18011.7	3674.54	4.902	1.53e-06	***
dt_33	1233.57	3449.85	0.3576	0.7209	
dt_34	-2783.05	3459.49	-0.8045	0.4217	
dt_35	-4714.95	3330.25	-1.416	0.1578	
dt_36	686.195	3357.91	0.2044	0.8382	

Střední hodnota závisle proměnné 40880.26
 Sm. odchylka závisle proměnné 14764.95
 Součet čtverců reziduí 1.68e+10
 Sm. chyba regrese 7322.275
 LSDV R-squared 0.785574
 Within R-squared 0.669925
 LSDV F(46, 313) 24.92847
 P-hodnota(F) 9.02e-80
 Logaritmus věrohodnosti -3689.159
 Akaiikovo kritérium 7472.318
 Schwarzovo kritérium 7654.965
 Hannan-Quinnovo kritérium 7544.942
 rho (koeficient autokorelace) 0.219760
 Durbin-Watsonova statistika 1.474215
 zde je poznámka o zkratkách statistik modelu

Joint test on named regressors -
 Testovací statistika: F(2, 313) = 26.6359
 s p-hodnotou = P(F(2, 313) > 26.6359) = 2.07711e-11

Test pro různé intercepty mezi skupinami -
 Nulová hypotéza: Skupiny mají společný intercept
 Testovací statistika: F(9, 313) = 11.0032
 s p-hodnotou = P(F(9, 313) > 11.0032) = 6.5639e-15

Wald joint test on time dummies -
 Nulová hypotéza: No time effects
 Asymptotická testovací statistika: Chí-kvadrát(35) = 209.62
 s p-hodnotou = 9.1011e-27

Obrázek 8 Model s fixními efekty se zahrnutím individuálních časových efektů 2

Zdroj: vlastní zpracování (software Gretl)

3.5.3 Model s náhodnými efekty

Obrázek 9 ukazuje výstup ze softwaru Gretl pro model s náhodnými efekty.

```
Model 9: Náhodné efekty (GLS), za použití 360 pozorování
Zahrnuto 10 průřezových jednotek
Délka časové řady = 36
Závisle proměnná: trzby

-----
                koeficient   směr. chyba   z       p-hodnota
-----
const           812.412      4629.81       0.1755   0.8607
navstevnost     44.3794       2.58726      17.15    5.97e-66 ***
zamestnanci     943.142       428.613       2.200    0.0278 **

Střední hodnota závisle proměnné   40880.26
Sm. odchylka závisle proměnné     14764.95
Součet čtverců reziduí              3.65e+10
Sm. chyba regrese                    10094.20
Logaritmus věrohodnosti              -3828.913
Akaikovo kritérium                   7663.826
Schwarzovo kritérium                 7675.485
Hannan-Quinnovo kritérium            7668.462
rho (koeficient autokorelace)        0.315936
Durbin-Watsonova statistika          1.308274
zde je poznámka o zkratkách statistik modelu

Rozptyl 'mezi' = 1.84664e+07
Rozptyl 'uvnitř' = 8.05191e+07
theta použitá pro částečné odstranění střední hodnoty = 0.671314
corr(y,yhat)^2 = 0.549862

Joint test on named regressors -
Asymptotická testovací statistika: Chí-kvadrát(2) = 305.785
s p-hodnotou = 3.97838e-67

Breusch-Paganův test -
Nulová hypotéza: Rozptyl chyb příslušejících jednotkám = 0
Asymptotická testovací statistika: Chí-kvadrát(1) = 136.066
s p-hodnotou = 1.93044e-31

Hausmanův test -
Nulová hypotéza: GLS odhady jsou konzistentní
Asymptotická testovací statistika: Chí-kvadrát(2) = 4.99785
s p-hodnotou = 0.0821733
```

Obrázek 9 Model s náhodnými efekty 2

Zdroj: vlastní zpracování (software Gretl)

Pro tento model jsou hodnoty informačních kritérií následující:

- Akaikovo kritérium – 7 663,826,
- Schwarzovo kritérium – 7 675,485,
- Hannan-Quinnovo kritérium – 7 668,462.

Hodnoty jsou nepatrně nižší než u modelu s náhodnými efekty bez zahrnutí proměnné „zaměstnanci“. P-hodnota proměnné „zaměstnanci“ se blíží hodnotě 0,05, proto není velmi významná. Proměnná „návštěvnost“ je velmi významná.

Hodnoty informačních kritérií jsou vyšší než u modelu s fixními efekty, proto lze říct, že model s fixními efekty je pro daná data vhodnější.

Model s náhodnými efekty bez zahrnutí individuálních časových efektů lze zapsat následovně:

$$y_{it} = 24587,6 + 25,4574x_{it1} + (\alpha + \varepsilon_i) + u_{it} \quad (3.7)$$

Obrázek 10 ukazuje výstup ze softwaru Gretl po použití modelu s náhodnými efekty a se zahrnutím časových individuálních efektů.

Z obrázku lze vyčíst, že obě vysvětlující proměnné jsou v modelu s náhodnými efekty se zahrnutím individuálních časových efektů velmi významné. Hodnoty časových efektů nemají klesající ani stoupající trend.

Tento model vyhodnocuje více časových efektů jako významné, než ukazoval stejný model bez zahrnutí proměnné „zaměstnanci“

Dle Hausmanova testu lze zamítnout nulovou hypotézu, že odhady jsou konzistentní. P-hodnota je vysoká. To znamená, že model s náhodnými efekty není lepší než model s fixními efekty.

Model 10: Náhodné efekty (GLS), za použití 360 pozorování
 Zahrnuto 10 průřezových jednotek
 Délka časové řady = 36
 Závisle proměnná: trzby

	koeficient	směr. chyba	z	p-hodnota	
const	13015.6	4969.78	2.619	0.0088	***
navstevnost	25.9961	3.20953	8.100	5.51e-16	***
zamestnanci	1322.92	472.939	2.797	0.0052	***
dt_2	3613.57	3281.69	1.101	0.2708	
dt_3	-765.989	3466.73	-0.2210	0.8251	
dt_4	-2487.92	3267.83	-0.7613	0.4465	
dt_5	-614.222	3364.72	-0.1825	0.8552	
dt_6	1887.54	3447.86	0.5475	0.5841	
dt_7	-2868.99	3271.85	-0.8769	0.3806	
dt_8	-13494.4	3303.85	-4.084	4.42e-05	***
dt_9	-12130.0	3294.81	-3.682	0.0002	***
dt_10	-10103.3	3341.76	-3.023	0.0025	***
dt_11	-13592.6	3367.77	-4.036	5.43e-05	***
dt_12	-7532.93	3425.72	-2.199	0.0279	**
dt_13	-8555.33	3351.79	-2.552	0.0107	**
dt_14	-9282.53	3332.28	-2.786	0.0053	***
dt_15	-6152.96	3324.60	-1.851	0.0642	*
dt_16	-7253.47	3333.69	-2.176	0.0296	**
dt_17	-10626.5	3284.70	-3.235	0.0012	***
dt_18	-1138.45	3341.26	-0.3407	0.7333	
dt_19	-2134.94	3315.42	-0.6439	0.5196	
dt_20	-1652.83	3406.59	-0.4852	0.6275	
dt_21	-507.777	3380.21	-0.1502	0.8806	
dt_22	-2404.04	3357.51	-0.7160	0.4740	
dt_23	-3786.78	3367.48	-1.125	0.2608	
dt_24	-4356.97	3362.59	-1.296	0.1951	
dt_25	-6322.21	3308.42	-1.911	0.0560	*
dt_26	-5788.53	3351.80	-1.727	0.0842	*
dt_27	-13934.2	3428.42	-4.064	4.82e-05	***
dt_28	3385.11	3356.78	1.008	0.3132	
dt_29	4874.80	3347.14	1.456	0.1453	
dt_30	12293.6	3560.22	3.453	0.0006	***
dt_31	1310.46	3383.11	0.3874	0.6985	
dt_32	17282.7	3540.86	4.881	1.06e-06	***
dt_33	801.829	3367.67	0.2381	0.8118	
dt_34	-3177.67	3373.82	-0.9419	0.3463	
dt_35	-4829.83	3303.21	-1.462	0.1437	
dt_36	426.011	3315.16	0.1285	0.8978	

Střední hodnota závisle proměnné 40880.26
 Sm. odchylka závisle proměnné 14764.95
 Součet čtverců rezidui 2.23e+10
 Sm. chyba regrese 8310.430
 Logaritmus věrohodnosti -3740.393
 Akaikovo kritérium 7556.785
 Schwarzovo kritérium 7704.457
 Hannan-Quinnovo kritérium 7615.502
 rho (koeficient autokorelace) 0.219760
 Durbin-Watsonova statistika 1.474215
 zde je poznámka o zkratkách statistik modelu

Rozptyl 'mezi' = 1.92137e+07
 Rozptyl 'uvnitř' = 5.36157e+07
 theta použitá pro částečné odstranění střední hodnoty = 0.731788
 corr(y,yhat)^2 = 0.716066

Joint test on named regressors -
 Asymptotická testovací statistika: Chí-kvadrát(2) = 82.8242
 s p-hodnotou = 1.03505e-18

Wald joint test on time dummies -
 Nulová hypotéza: No time effects
 Asymptotická testovací statistika: Chí-kvadrát(35) = 215.114
 s p-hodnotou = 8.9032e-28

Breusch-Paganův test -
 Nulová hypotéza: Rozptyl chyb příslušejících jednotkám = 0
 Asymptotická testovací statistika: Chí-kvadrát(1) = 291.37
 s p-hodnotou = 2.5009e-65

Hausmanův test -
 Nulová hypotéza: GLS odhady jsou konzistentní
 Asymptotická testovací statistika: Chí-kvadrát(2) = 0.67444
 s p-hodnotou = 0.713752

Obrázek 10 Model s náhodnými efekty se zahrnutím individuálních časových efektů 2

Zdroj: vlastní zpracování (software Gretl)

3.6 Výběr optimálního modelu

Nejkvalitnější model panelových dat, kde jako závislá proměnná byly tržby a nezávislé proměnné byly návštěvnost a počet zaměstnanců, bude určen opět pomocí součtu čtverců reziduí (viz Tabulka 18).

Tabulka 18 Součty čtverců reziduí 2

model	součet čtverců reziduí
6 – Spojený regresní model	35 000 000 000
7 – Model s fixními efekty	28 000 000 000
8 – Model s fixními efekty se zahrnutím individuálních časových efektů	16 800 000 000
9 – Model s náhodnými efekty	36 500 000 000
10 – Model s náhodnými efekty se zahrnutím individuálních časových efektů	22 300 000 000

Zdroj: vlastní zpracování

Nejnižší součet čtverců reziduí má model s fixními efekty se zahrnutím individuálních časových efektů. Stejně jako v případě vyhodnocení modelů bez zahrnutí proměnné „zaměstnanci“:

- Hausmanův test potvrdil, že je pro daná data vhodnější model fixních efektů než model s náhodnými efekty,
- byla zamítnuta hypotéza o nepotřebnosti časových individuálních efektů, a tedy je tento model vhodnější než model fixních efektů bez zahrnutí časových individuálních proměnných.

3.6.1 Analýza výsledků vybraného modelu

Po zahrnutí vysvětlující proměnné počtu zaměstnanců vyšel jako nejvhodnější model pro daná data model s fixními efekty se zahrnutím individuálních časových efektů.

Model lze zapsat takto:

$$y_{it} = 15298,1 + 24,9235x_{it1} + 1131,03x_{it2} + \varepsilon_i + \lambda_t + u_{it} \quad (3.8)$$

kde:

ε_i vyjadřuje individuální efekty stánků,

λ_t vyjadřuje individuální časové efekty.

Z rovnice lze vyčíst, že pokud se návštěvnost zvýší o jednoho člověka, potom tržba stánku vzroste v průměru o 24,9235 Kč. Pokud počet lidí, kteří ve stánku obsluhují, vzroste o jednoho zaměstnance, tržba stánku se v průměru zvýší o 1 131,03 Kč. Pokud by vliv všech nezávislých proměnných v modelu byl nulový, bude tržba stánku rovna 15 298,1 Kč.

Vliv proměnné „návštěvnost“ je velmi významný, nicméně vliv proměnné „zaměstnanci“ na model nemá až tak významný vliv.

Tabulka 19 shrnuje individuální časové efekty a jejich významnost.

Tabulka 19 Koeficienty individuálních časových efektů 2

zápas	koeficient časového efektu	významnost	zápas	koeficient časového efektu	významnost
1	–		19	-1860,49	
2	3715,46		20	-1130,19	
3	-378,903		21	-22,9448	
4	-2511,51		22	-1969,38	
5	-345,868		23	-3356,21	
6	2255,32		24	-3978,75	
7	-2812,58		25	-6058,59	*
8	-13325,6	***	26	-5445,48	
9	-11952,1	***	27	-13827,5	***
10	-9785,08	***	28	3793,91	
11	-13286,4	***	29	5264,52	
12	-6945,79	*	30	13053,9	***
13	-8211,85	**	31	1804,41	
14	-8971,09	***	32	18011,7	***
15	-5840,34		33	1233,57	
16	-6982,58		34	-2783,05	
17	-10558		35	-4714,95	
18	-916,471		36	686,195	

Zdroj: vlastní zpracování

Počet významných časových efektů je o jeden nižší než u stejného modelu bez zahrnutí vysvětlující proměnné „zaměstnanci“. V tomto případě jsou velmi významné časové efekty u zápasů 8, 9, 10, 11, 14, 27, 30 a 32.

V následujících tabulkách jsou opět seskupeny zápasy dle třech kritérií a skupiny jsou porovnány s hodnotami časových koeficientů.

Tabulka 20 Hodnoty individuálních časových efektů v rámci zápasů se stejným soupeřem 2

skupiny	hodnoty koeficientů časových efektů			
2, 16, 26,	3715,46	-6982,58	-5445,48	-4714,95
12, 18, 22	2255,32	13053,9		
6, 30	-6945,79	-916,471	-1969,38	
5, 20, 36	-345,868	-1130,19	686,195	
8, 19	-13325,6	-1860,49		
10, 31	-9785,08	1804,41		
14, 29	-8971,09	5264,52		
13, 24	-8211,85	-3978,75		
4, 28, 34	-2511,51	3793,91	-2783,05	
15, 33	-5840,34	1233,57		
9, 27	-11952,1	-13827,5		
11, 17, 21	-13286,4	-10558	-22,9448	
3, 32	-378,903	18011,7		

Zdroj: vlastní zpracování

V Tabulka 20 jsou zápasy seskupeny podle stejného soupeře. Hodnoty časových koeficientů jsou v rámci skupin velmi rozdílné. Nelze předpokládat stejný časový efekt u zápasů se stejným soupeřem.

V Tabulka 21 jsou zápasy rozděleny podle času začátku utkání. Ani v tomto případě nejsou časové efekty v rámci jednotlivých skupin podobné. Nelze tedy předpokládat, že by měl stejný čas zápasu stejný efekt.

Tabulka 21 Hodnoty individuálních časových efektů v rámci zápasů se stejným časem začátku utkání 2

	čas utkání			
	15:00	17:00	18:00	19:00
hodnoty koeficientů časových efektů	-13325,6	3715,46	2255,32	-378,903
	-11952,1	-2812,58	-9785,08	-2511,51
	-13286,4	686,195	-6945,79	-345,868
	-5840,34		-6982,58	-8211,85
	-10558		-1860,49	-8971,09
	-1130,19		-3978,75	-916,471
	-22,9448		3793,91	
	-1969,38		5264,52	
	-3356,21		13053,9	
	-6058,59		18011,7	
	-5445,48			
	-13827,5			
	1804,41			
	1233,57			
	-2783,05			
	-4714,95			

Zdroj: vlastní zpracování

V Tabulka 22 jsou zápasy seskupeny podle posledního kritéria, a to podle dne konání utkání. Ani z této tabulky nelze usuzovat, že by měl konkrétní den konání zápasu podobný časový efekt.

Tabulka 22 Hodnoty individuálních časových efektů v rámci zápasů se stejným dnem konání 2

	den zápasu		
	středa	sobota	neděle
hodnoty koeficientů časových efektů	3715,46	-2511,51	-378,903
	-2812,58	-345,868	-11952,1
		2255,32	-6945,79
		-13325,6	-8971,09
		-9785,08	-5840,34
		-13286,4	-1860,49
		-8211,85	-1130,19
		-6982,58	-22,9448
		-10558	-1969,38
		3793,91	-3356,21
		5264,52	-3978,75
		1804,41	-6058,59
		-4714,95	-5445,48
		686,195	-13827,5
			13053,9
			18011,7
			1233,57
			-2783,05

Zdroj: vlastní zpracování

V Tabulka 23 jsou uvedeny koeficienty vztahené k jednotkám. Ty ukazují hodnotu tržeb jednotlivých stánků, kdyby návštěvnost ani počet zaměstnanců neměly na výši tržeb žádný vliv.

Tabulka 23 Konstanty vztahené k jednotkám 2

číslo stánku	konstanta
stánek 1	11 522,83
stánek 2	18 620,37
stánek 3	18 046,17
stánek 4	11 726,20
stánek 5	8 055,35
stánek 6	20 426,40
stánek 7	12 691,23
stánek 8	18 497,24
stánek 9	19 762,93

číslo stánku	konstanta
stánek 10	13 632,03

Zdroj: vlastní zpracování

Stejně jako u modelu bez zahrnutí počtu zaměstnanců by měl nejvyšší tržby stánek 6, následovaly by stánky 9, 2, 8 a 3. Nejnižší tržbu by měl stánek 7.

Při porovnání konstant s průměrnou tržbou (viz Tabulka 24) lze říct, že velký vliv návštěvnosti a počtu zaměstnanců na výši tržeb je u stánku 2, 3 a 4. Čím vyšší je návštěvnost a počet zaměstnanců, tím vyšší jsou tržby. Naopak nejmenší vliv je, stejně jako u modelu bez počtu zaměstnanců, u stánku 7.

Tabulka 24 Porovnání průměrných tržeb a konstant 2

	průměrné tržby	konstanty vztažené k jednotkám
stánek 1	34434,6	11522,83
stánek 2	54493,6	18620,37
stánek 3	50770,4	18 046,17
stánek 4	43302,1	11726,2
stánek 5	33167,1	8055,35
stánek 6	48940,3	20426,4
stánek 7	25505,1	12691,23
stánek 8	41870,9	18497,24
stánek 9	43379,8	19762,93
stánek 10	32938,7	13632,03

Zdroj: vlastní zpracování

3.7 Závěrečné porovnání a vyhodnocení

V předchozích podkapitolách byly zpracovány jednotlivé modely panelových dat, které zkoumaly závislost tržeb nejprve pouze na návštěvnosti. Poté byly zpracovány modely, které zkoumaly závislost tržeb na návštěvnosti a počtu zaměstnanců jednotlivých stánků.

V obou případech vyšel jako nejvhodnější pro daná data model s fixními efekty se zahrnutím časových individuálních efektů.

Zápasy byly seskupeny podle těchto tří kritérií:

- stejný soupeř,
- stejný čas začátku utkání,

- stejný den v týdnu, kdy se utkání konalo.

V rámci těchto skupin byly porovnány hodnoty koeficientů individuálních časových efektů. Ani v jednom porovnání nebyly časové individuální efekty v rámci skupin podobné. Pomocí hodnot koeficientů individuálních časových efektů tedy nelze vyvodit žádné důsledky.

Následně byly vyhodnoceny konstanty vztažené k jednotkám, které ukazují hodnotu tržeb jednotlivých stánků, kdyby žádná vysvětlující proměnná neměla na výši tržeb vliv. Pokud by návštěvnost ani počet zaměstnanců stánku neměly na výši tržby vliv, nejvyšší tržbu by měl stánek 6 a nejnižší stánek 7.

U modelu, do kterého byla zahrnuta proměnná „zaměstnanci“, vyšla p-hodnota této proměnné vyšší než 0,05. Počet zaměstnanců, na rozdíl od návštěvnosti, je v modelu málo významný. Počet zaměstnanců tedy nemá tak velký vliv na výši tržeb jako návštěvnost.

Tabulka 25 uvádí rozdíly mezi průměrnou tržbou stánků a tržbou, které by dosahovaly, pokud by jednotlivé vysvětlující proměnné v modelech neměly na její výši vliv. Pokud by návštěvnost a počet zaměstnanců stánku neměly vliv na výši tržeb, nejméně by se tržby změnily u stánku 7. Naopak pokud se mění návštěvnost a počet zaměstnanců, nejvíce se tržby mění u stánků 2, 3 a 4.

Tabulka 25 Rozdíly mezi průměrnou tržbou a konstantami

	rozdíly mezi průměrnou tržbou a tržbou bez působení vysvětlujících proměnných	
	návštěvnost	návštěvnost, počet zaměstnanců
stánek 1	13552,25	22911,77
stánek 2	26727,18	35873,23
stánek 3	23286,09	32724,23
stánek 4	21920,62	31575,9
stánek 5	14130,64	25111,75
stánek 6	14974,25	28513,9
stánek 7	8853,88	12813,87
stánek 8	14344,93	23373,66
stánek 9	14400,81	23616,87
stánek 10	10736,36	19306,67

Zdroj: vlastní zpracování

Pokud porovnáme dva výsledné modely s fixními efekty se zahrnutím individuálních časových efektů mezi sebou, jejich výsledky jsou podobné. Významnost vysvětlující proměnné „zaměstnanci“ je malá, proto není interpretace modelů příliš rozdílná.

3.8 Doporučení

Na základě dat, které jsou v této práci analyzována, lze potvrdit, že na výši tržeb stánků má výrazný vliv jejich návštěvnost. Návštěvnost stánků závisí na návštěvnosti jednotlivých sektorů, a tedy návštěvnosti celého stadionu. Stánky zákazníci navštěvují během fotbalového zápasu nerovnoměrně. Před zápasem a během přestávky je návštěvnost výrazně vyšší než v průběhu utkání. Je proto důležité za krátký čas obsloužit co nejvíce zákazníků a snížit jejich dobu čekání. Provozovatel stánků s občerstvením by mohl zavést možnost platit pouze bezhotovostně. Zkrátila by se tak doba transakce, a to by přispělo k vyšší návštěvnosti stánku a tím i k vyšším tržbám.

Provozovatel by se tak měl soustředit na zvyšování návštěvnosti především u stánků 2, 3 a 4, protože u nich návštěvnost ovlivňuje tržby nejvíce ze všech stánků.

S rostoucím počtem zaměstnanců, stoupají tržby. Nicméně vliv počtu zaměstnanců, kteří stánek obsluhují, je málo významný. Proto není podstatné počty zaměstnanců zvyšovat. I vzhledem k tomu, že jsou stánky uzpůsobené pro omezený počet zaměstnanců.

Nelze předpokládat podobné časové efekty u zápasů podle stejného soupeře, času začátku utkání nebo dne v týdnu, kdy se zápas hraje. Provozovatel stánků s občerstvením tak nemůže podle těchto kritérií vyvozovat žádné důsledky.

ZÁVĚR

Tato diplomová práce řešila problematiku analýzy panelových dat v podnikové praxi. Jejím cílem bylo aplikovat analýzu panelových dat na konkrétní problém z podnikové praxe, vybrat vhodný model pro daná data a odhadnout jeho parametry. Následně odhadnuté parametry interpretovat a navrhnout doporučení. Práce je rozdělena do tří kapitol, které obsahují podkapitoly.

První kapitola se zaměřuje na teoretické vymezení panelových dat. Stručně popisuje jejich vývoj. Popisuje charakteristiky panelových dat a možnosti, kterými lze data zapisovat. Dále uvádí výhody a nevýhody, kterými se panelová data, na rozdíl od dat průřezových nebo časových, vyznačují. Tato kapitola dále představuje oblasti a možnosti využití panelových dat v praxi. V poslední části je popsán panelový výzkum a jsou uvedeny některé konkrétní panelové výzkumy ze světa i z České republiky.

Druhá kapitola rozebírá základní modely, kterými lze panelová data analyzovat. Nejprve je popsán spojený regresní model. Poté jsou rozebrány modely s individuálními efekty, konkrétně jsou popsány model s fixními efekty a model s náhodnými efekty. Dále se kapitola věnuje modelům, které zahrnují zpožděnou závisle proměnnou, a to dynamickému a autoregresnímu modelu.

Třetí kapitola se zaměřuje na splnění cíle této diplomové práce. V této kapitole byla analyzována data tržeb, návštěvnosti a počtu zaměstnanců stánků s občerstvením na fotbalovém stadionu. Nejprve byly uvedeny popisné statistiky jednotlivých proměnných. Následoval odhad parametrů vybraných modelů a interpretace odhadnutých parametrů. Model, který byl vybrán jako nejvhodnější pro daná data, byl následně podrobněji popsán. Byla zkoumána závislost tržeb na návštěvnosti a počtu zaměstnanců jednotlivých stánků. Nejprve byla do modelu zahrnuta jako vysvětlující proměnná pouze návštěvnost, poté byl přidán i počet zaměstnanců a analýza byla provedena znovu.

Po vyhodnocení optimálních modelů bylo zjištěno, že návštěvnost stánku ovlivňuje tržby výrazněji než počet zaměstnanců, kteří stánek obsluhují. Podle stejného soupeře, stejného času začátku utkání ani podle stejného dne v týdnu, kdy se zápas hraje, nelze předpokládat stejný časový efekt. Dále bylo zjištěno, u kterých stánků mají změny v návštěvnosti a počtu

zaměstnanců největší vliv na tržby a u kterých naopak nejmenší. Na základě těchto faktů bylo v závěrečné části této práce navrženo doporučení.

Tato diplomová práce pomáhá s orientací mezi modely panelových dat a především se zaměřuje na jejich aplikaci v rámci podnikové praxe.

POUŽITÁ LITERATURA

Tištěné zdroje:

- [1] ANDREß, Hans-Jurgen; GOLSCH, Katrin a W.SCHMIDT, Alexander. *Applied Panel Data Analysis for Economic and Social Surveys*. Springer, 2013. ISBN 987-3-642-32913-5.
- [2] ARLT, Josef a ARLTOVÁ, Markéta. *Ekonomické časové řady*. Praha: Professional Publishing, 2009. ISBN 978-80-86946-85-6.
- [3] BALTAGI, Badi H. *Econometric Analysis of Panel Data*. 6th ed. Cham: Springer Cham, 2021. ISBN 9783030539528.
- [4] BOND, Stephen R. Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal*. 2002, s. 141–162.
- [5] CIPRA, Tomáš. *Finanční ekonometrie*. 2., upr. vyd. Praha: Ekopress, 2013. ISBN 978-80-86929-93-4.
- [6] GREENE, William H. *Econometric Analysis*. 5th ed. Pearson Education, 2003. ISBN 0-13-066189-9.
- [7] GUJARATI, Damodar N. a PORTER, Dawn C. *Basic econometrics*. 5th ed. McGraw-Hill/Irwin, 2009. ISBN 978-0-07-337577-9.
- [8] HEBÁK, Petr; HUSTOPECKÝ, Jiří; JAROŠOVÁ, Eva a MALÁ, Ivana. *Vícerozměrné statistické metody*. Praha: Informatorium, 2005. ISBN 80-7333-025-3.
- [9] HSIAO, Cheng. *Analysis of Panel Data*. 4th ed. Cambridge University Press, 2022. ISBN 978-1-316-51210-4.
- [10] KOČENDA, Evžen a ČERNÝ, Alexandr. *Elements of time series econometrics: an applied approach*. Third edition. Prague: Charles University in Prague, Karolinum Press, 2015. ISBN 978-80-246-3199-8.
- [11] KUBANOVÁ, Jana. *Statistické metody pro ekonomickou a technickou praxi*. 3. dopl. vyd. Statisticky, 2008. ISBN 978-80-85659-47-4.
- [12] NERLOVE, Mark. *Essays in Panel Data Econometrics*. Cambridge University Press, 2002. ISBN 978-0-521-815345-5.
- [13] TAHAL, Radek. *Marketingový výzkum: postupy, metody, trendy*. 2. vydání. *Expert (Grada)*. Praha: Grada Publishing, 2022. ISBN 978-80-271-3535-6.
- [14] WOOLDRIDGE, Jeffrey M. *Introductory Econometrics: A Modern Approach*. 3rd ed. Mason: Thomson South-Western, 2006. ISBN 0-324-32348-4.

Elektronické zdroje:

1. *About POLPAN*. Online. POLPAN. b.r. Dostupné z: <https://polpan.org/en/about-polpan/>. [cit. 2024-06-23].
2. *About us*. Online. 29th International Panel Data Conference. 2024. Dostupné z: <https://ipdc2024.org>. [cit. 2024-06-17].
3. *Cohorts*. Online. National Longitudinal Surveys. b.r. Dostupné z: <https://www.nlsinfo.org/content/cohorts>. [cit. 2024-05-23].
4. *Detailed Explanation of Panel Data*. Online. Medium. 2022. Dostupné z: <https://medium.com/geekculture/detailed-explanation-of-panel-data-how-to-identify-balanced-and-unbalanced-panel-data-fd973fa788ae>. [cit. 2024-06-02].
5. *Dotazování v 6. ročníku bylo úspěšně ukončeno!*. Online. CHPS. 2024. Dostupné z: <https://chps.soc.cas.cz/cs/aktuality/dotazovani-v-6-rocniku-bylo-uspesne-ukonceno>. [cit. 2024-05-23].
6. HOCH, Inrving. Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data. Online. 1962. Dostupné z: <https://www.jstor.org/stable/1911286>. [cit. 2024-06-17].
7. *Inform your marketing strategies with panel data*. Online. Kantar. 2023. Dostupné z: <https://www.kantar.com/north-america/inspiration/research-services/inform-your-marketing-strategies-with-panel-data-pf>. [cit. 2024-06-23].
8. *Introduction to the Fundamentals of Panel Data*. Online. Aptech. 2019. Dostupné z: <https://www.aptech.com/blog/introduction-to-the-fundamentals-of-panel-data/>. [cit. 2024-06-02].
9. *Jak na GPG z pohledu makroekonomie*. Online. Ministerstvo práce a sociálních věcí, 2020. ISBN 978-80-7421-239-0. Dostupné z: https://www.rovnaodmena.cz/www/img/uploads/Makroanaliza_07_22022021.pdf. [cit. 2024-06-23].
10. *Národní výzkumy*. Online. PIAAC. c2011–2024. Dostupné z: <https://piaac.cz/o-vyzkumu/navazne-vyzkumy/>. [cit. 2024-05-29].
11. LIPPS, Oliver. Panel Surveys. Online. *FORS Guide*. Č. no.14. Dostupné z: https://forscenter.ch/wp-content/uploads/2021/03/forsguide_panel_surveys_lipps_v_1.pdf. [cit. 2024-06-25].
12. *Panel Study of Income Dynamics*. Online. PSID. c2024. Dostupné z: <https://psidonline.isr.umich.edu>. [cit. 2024-05-23].
13. *Představení výzkumu*. Online. CHPS. c2023. Dostupné z: <https://chps.soc.cas.cz/cs/predstaveni-vyzkumu>. [cit. 2024-05-23].

14. *Regression using Panel data for beginners*. Online. Medium. 2023. Dostupné z: <https://medium.com/@laibamehnaz/regression-using-panel-data-for-beginners-f9b636deb459>. [cit. 2024-06-23].
15. *The Pooled OLS Model*. Online. Tilburg Science Hub. B.r. Dostupné z: <https://tilburgsciencehub.com/topics/analyze/causal-inference/panel-data/pooled-ols/>. [cit. 2024-06-23].
16. *The Random Effects Model*. Online. Tilburg Science Hub. b.r. Dostupné z: <https://tilburgsciencehub.com/topics/analyze/causal-inference/panel-data/random-effects/>. [cit. 2024-06-23].
17. *The Role of Panel Research in Enhancing Patient Experience in the Healthcare Industry*. Online. Data Spring. 2023. Dostupné z: <https://www.d8aspring.com/blog/the-role-of-panel-research-in-the-healthcare-industry>. [cit. 2024-06-23].
18. *The Survey of Adult Skills (PIAAC)*. Online. OECD. b.r. Dostupné z: <https://www.oecd.org/skills/piaac/about/#d.en.481111>. [cit. 2024-05-29].
19. *The wide and long data format for repeated measures data*. Online. The Analysis Factor. b.r. Dostupné z: <https://www.theanalysisfactor.com/wide-and-long-data/>. [cit. 2024-06-17].

SEZNAM PŘÍLOH

Příloha 1	Výchozí data	73
Příloha 2	Vývoj tržeb	76
Příloha 3	Vývoj návštěvnosti	77
Příloha 4	Vývoj počtu zaměstnanců	78

Příloha 1 Výchozí data

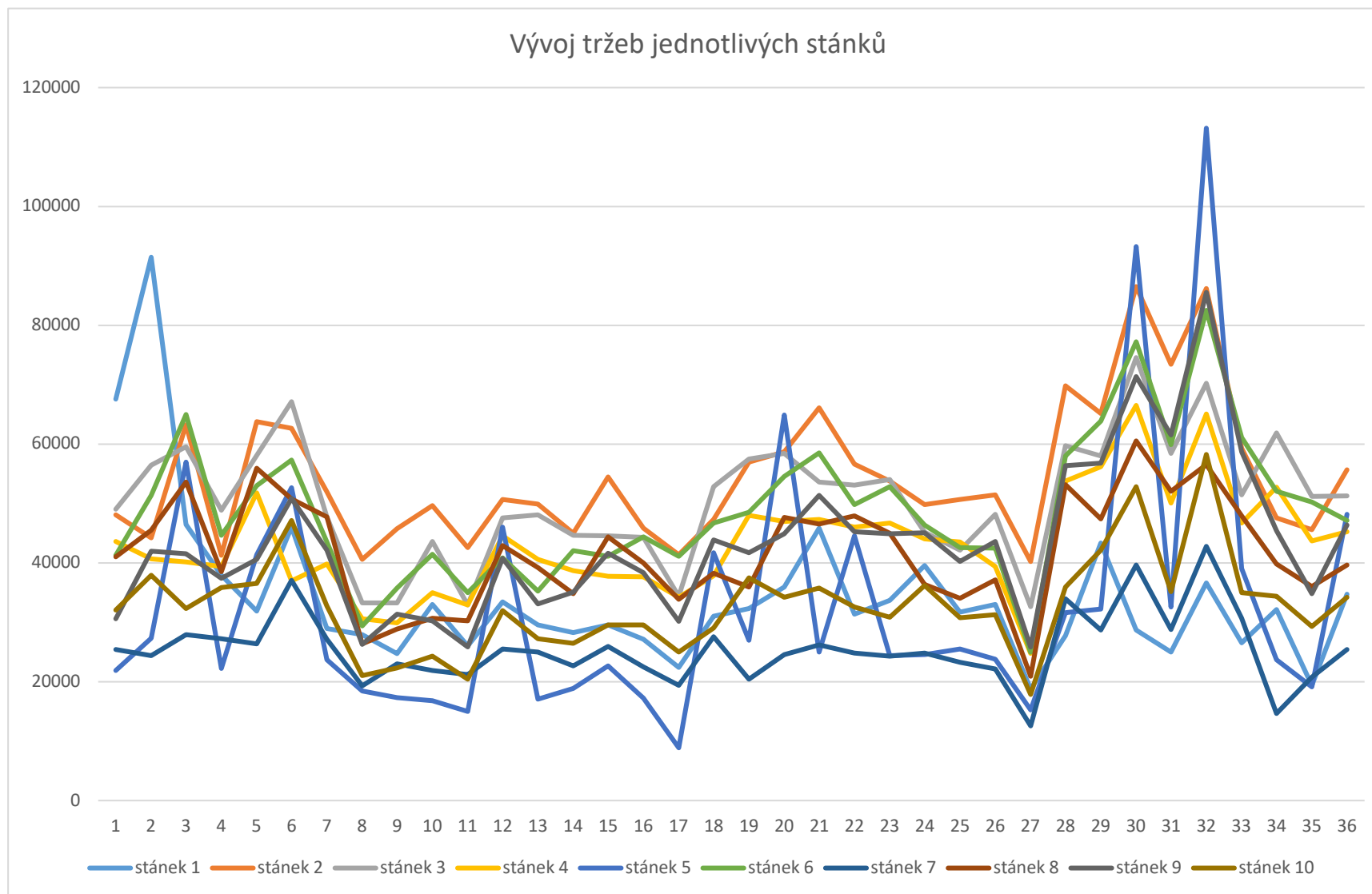
stánek 1				stánek 2				stánek 3				stánek 4			
zápas	tržby (v Kč)	návštěvno st	počet zaměstna nců	zápas	tržby (v Kč)	návštěvno st	počet zaměstna nců	zápas	tržby (v Kč)	návštěvno st	počet zaměstna nců	zápas	tržby (v Kč)	návštěvno st	počet zaměstna nců
t	Y_{it}	X_{it1}	X_{it2}	t	Y_{it}	X_{it1}	X_{it2}	t	Y_{it}	X_{it1}	X_{it2}	t	Y_{it}	X_{it1}	X_{it2}
1	67536	588	8	1	48080	1178	10	1	49036	1010	9	1	43576	811	9
2	91440	611	9	2	44184	1071	9	2	56484	1024	9	2	40720	968	9
3	46476	645	9	3	63228	1016	9	3	59520	1021	9	3	40136	1297	9
4	37920	607	9	4	41304	1174	9	4	48828	1044	9	4	39360	752	9
5	31892	571	9	5	63764	900	9	5	58084	904	9	5	51776	1149	9
6	45980	633	9	6	62672	997	9	6	67176	1002	9	6	36936	1273	9
7	28920	598	9	7	51932	1107	9	7	47776	958	9	7	39800	855	9
8	27956	593	10	8	40596	1094	10	8	33240	893	10	8	30616	768	10
9	24724	563	10	9	45804	1067	9	9	33244	876	10	9	29932	810	10
10	33020	625	10	10	49600	1206	10	10	43584	1007	10	10	35004	837	10
11	25988	532	10	11	42576	1047	10	11	32900	839	11	11	32940	804	11
12	33432	543	10	12	50688	855	9	12	47580	859	10	12	44564	1092	12
13	29580	602	9	13	49896	1236	10	13	48116	1015	10	13	40556	837	11
14	28232	560	10	14	44972	1138	9	14	44664	987	9	14	38736	847	11
15	29580	573	10	15	54428	1139	10	15	44516	963	10	15	37732	850	11
16	27172	553	10	16	45880	1136	10	16	44268	953	10	16	37672	776	11
17	22376	514	9	17	41376	1083	10	17	34384	963	9	17	34172	766	10
18	31048	514	10	18	47356	1083	10	18	52844	963	11	18	37848	766	11
19	32280	586	10	19	56976	1275	10	19	57448	1049	10	19	47968	866	11
20	35912	586	9	20	58692	923	10	20	58436	928	10	20	46940	1179	10
21	45860	731	10	21	66096	1242	10	21	53640	1078	11	21	47324	953	10
22	31340	616	10	22	56584	1239	10	22	53052	1048	10	22	46056	932	10
23	33724	596	10	23	53736	1192	10	23	53996	1024	11	23	46716	911	10
24	39580	593	10	24	49824	1113	10	24	45124	966	10	24	44016	865	10
25	31736	593	9	25	50688	1133	9	25	42184	949	10	25	43552	912	10
26	32988	581	10	26	51460	1121	10	26	48140	904	10	26	39388	853	10
27	18764	426	10	27	40228	910	10	27	32628	690	11	27	24764	546	10
28	27724	628	10	28	69784	1112	10	28	59736	954	10	28	53816	932	10
29	43312	604	10	29	65192	1177	10	29	58032	958	10	29	56232	899	10
30	28728	674	10	30	86512	1062	10	30	74556	1067	10	30	66508	1356	10
31	24980	650	10	31	73384	1307	10	31	58432	1048	11	31	50052	972	10
32	36592	667	10	32	86148	1050	10	32	70220	1055	10	32	65052	1341	10
33	26568	597	10	33	59304	1180	10	33	51480	955	11	33	46732	879	10
34	32168	602	10	34	47588	1164	10	34	61912	1078	11	34	52724	851	11
35	19396	477	10	35	45548	1011	10	35	51212	989	10	35	43712	773	10
36	34720	608	10	36	55688	1133	10	36	51264	984	8	36	45248	796	10

stánek 5				stánek 6				stánek 7				stánek 8			
zápas	tržby (v Kč)	návštěvno st	počet zaměstna nců	zápas	tržby (v Kč)	návštěvno st	počet zaměstna nců	zápas	tržby (v Kč)	návštěvno st	počet zaměstna nců	zápas	tržby (v Kč)	návštěvno st	počet zaměstna nců
t	Y _{it}	X _{it1}	X _{it2}	t	Y _{it}	X _{it1}	X _{it2}	t	Y _{it}	X _{it1}	X _{it2}	t	Y _{it}	X _{it1}	X _{it2}
1	21892	446	9	1	41292	412	11	1	25436	299	5	1	41008	483	8
2	27296	504	9	2	51356	733	11	2	24420	383	5	2	45508	625	8
3	56980	1265	9	3	64984	1079	11	3	27920	690	5	3	53580	976	8
4	22208	458	9	4	44640	427	11	4	27192	291	4	4	38500	462	8
5	41388	1120	9	5	53016	956	11	5	26408	611	5	5	55892	864	8
6	52668	1241	9	6	57276	1059	11	6	37092	677	5	6	50724	958	8
7	23740	491	9	7	43480	531	11	7	27124	392	5	7	47780	623	8
8	18440	449	10	8	29368	440	13	8	19340	318	5	8	26336	478	9
9	17336	456	11	9	35724	496	13	9	23004	359	4	9	28912	547	9
10	16828	460	11	10	41456	508	15	10	21912	332	5	10	30688	495	9
11	15000	431	11	11	35016	461	15	11	21224	345	5	11	30276	529	9
12	46004	1064	12	12	40820	908	15	12	25480	581	5	12	42948	821	10
13	17068	447	11	13	35280	500	15	13	25008	350	5	13	39188	511	10
14	18844	449	12	14	42096	528	13	14	22660	357	5	14	34860	580	10
15	22636	497	11	15	41128	634	13	15	25968	350	4	15	44372	619	10
16	17216	414	11	16	44364	498	13	16	22536	317	4	16	39988	524	10
17	8876	367	11	17	41112	401	10	17	19428	284	5	17	33908	455	8
18	41708	367	11	18	46720	401	12	18	27616	284	5	18	38244	455	10
19	26996	542	11	19	48536	454	12	19	20456	319	5	19	35896	492	10
20	64916	1149	12	20	54520	981	13	20	24528	627	5	20	47616	887	10
21	25000	521	12	21	58524	760	15	21	26224	393	5	21	46556	650	10
22	44560	617	12	22	49780	745	14	22	24808	401	5	22	47884	662	10
23	24352	496	12	23	52836	741	15	23	24276	362	5	23	45028	618	10
24	24568	475	13	24	46396	550	15	24	24836	368	5	24	36360	566	10
25	25488	470	10	25	42616	537	15	25	23304	343	5	25	34032	585	9
26	23804	463	12	26	42508	530	15	26	22156	362	5	26	37172	572	10
27	15252	299	12	27	24960	268	15	27	12564	206	5	27	20916	292	10
28	31664	513	12	28	58016	713	15	28	33928	417	5	28	53196	619	10
29	32216	519	12	29	63832	725	14	29	28660	396	5	29	47416	627	10
30	93236	1322	13	30	77224	1128	15	30	39668	721	5	30	60532	1020	10
31	32596	538	12	31	59868	808	15	31	28752	425	5	31	52052	645	10
32	113168	1307	13	32	82520	1115	15	32	42784	713	4	32	56520	1009	10
33	39024	523	12	33	61140	746	15	33	30708	379	5	33	47956	635	10
34	23744	452	12	34	52036	555	15	34	14616	318	5	34	39852	530	10
35	19112	381	10	35	50276	422	12	35	20728	282	4	35	35996	446	9
36	48192	545	13	36	47136	501	13	36	25420	344	4	36	39660	501	10

stánek 9				stánek 10			
zápas	tržby (v Kč)	návštěvnost	počet zaměstnanců	zápas	tržby (v Kč)	návštěvnost	počet zaměstnanců
t	V _{it}	X _{it1}	X _{it2}	t	V _{it}	X _{it1}	X _{it2}
1	30576	465	8	1	32072	382	8
2	41996	609	8	2	37932	496	8
3	41540	990	8	3	32300	704	8
4	37424	439	8	4	35860	379	8
5	40624	877	8	5	36560	624	8
6	50856	972	8	6	47140	691	8
7	42172	570	8	7	32764	475	8
8	26260	468	9	8	21048	358	9
9	31360	537	9	9	22344	412	9
10	30284	491	10	10	24296	396	10
11	25872	506	10	11	20440	394	10
12	40800	833	11	12	31936	593	10
13	33068	507	10	13	27196	409	10
14	35072	577	10	14	26436	450	10
15	41604	585	10	15	29520	453	9
16	38372	513	10	16	29568	411	10
17	30200	437	10	17	25036	369	9
18	43864	437	10	18	29084	369	9
19	41708	523	10	19	37476	380	8
20	44900	900	10	20	34256	640	8
21	51368	698	10	21	35772	527	9
22	45232	671	10	22	32600	512	9
23	44872	626	10	23	30884	481	9
24	45108	580	10	24	36220	483	9
25	40244	586	10	25	30728	456	9
26	43620	587	10	26	31284	437	9
27	25836	310	9	27	17856	259	9
28	56364	644	10	28	35964	491	9
29	56784	623	10	29	42116	496	9
30	71380	1035	10	30	52792	736	9
31	61516	704	10	31	35124	541	9
32	85532	1024	10	32	58256	728	9
33	58708	659	10	33	34972	505	9
34	45404	537	10	34	34428	446	9
35	34820	411	10	35	29300	343	9
36	46332	509	10	36	34232	432	9

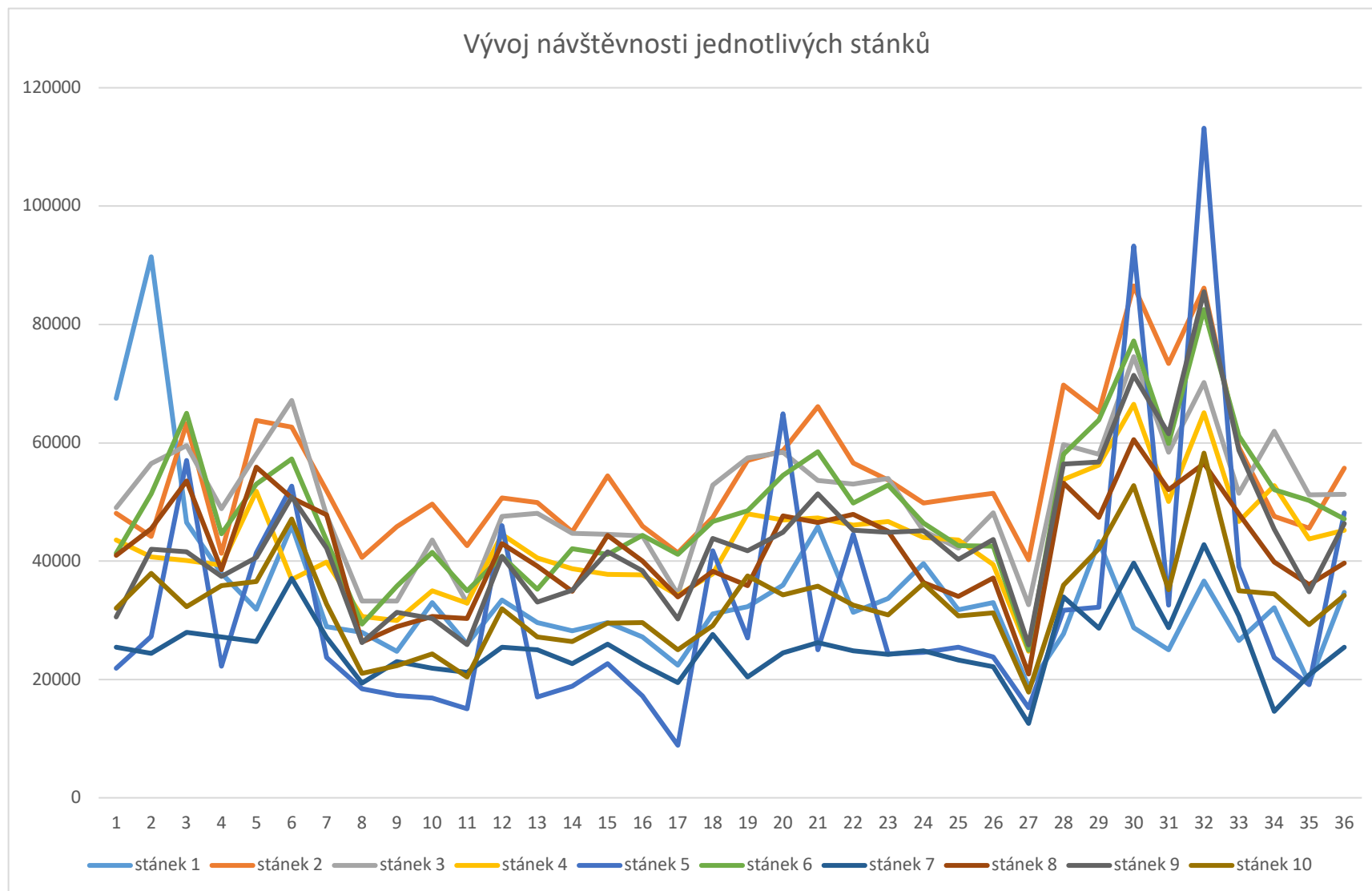
Příloha 2 Vývoj tržeb

76



Příloha 3 Vývoj návštěvnosti

77



Příloha 4 Vývoj počtu zaměstnanců

78

