

Univerzita Karlova v Praze  
Matematicko-fyzikální fakulta

## DIPLOMOVÁ PRÁCE



Radek Hendrych

### **Ekonometrické soustavy simultánních rovnic v životním pojištění**

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: prof. RNDr. Tomáš Cipra, DrSc.

Studijní program: Matematika

Studijní obor: Pravděpodobnost, matematická statistika a ekonometrie

Studijní plán: Ekonometrie

2010

Na tomto místě bych rád poděkoval vedoucímu mé diplomové práce prof. RNDr. Tomáši Ciprovi, DrSc. za cenné rady a připomínky, které mi poskytoval v průběhu přípravy tohoto textu a během společných konzultací.

Prohlašuji, že jsem svou diplomovou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 15. dubna 2010.

Radek Hendrych

# Obsah

Úvod	6
<b>1 Ekonometrické soustavy simultánních rovnic</b>	<b>7</b>
1.1 Formulace soustavy simultánních rovnic . . . . .	9
1.2 Problematika identifikace . . . . .	12
1.3 Nekonzistence OLS-odhadu v rámci simultánního modelování . . . . .	17
1.4 Odhady parametrů v rámci soustav simultánních rovnic . . . . .	19
1.4.1 Nepřímý odhad metodou nejmenších čtverců . . . . .	19
1.4.2 Odhad metodou instrumentálních proměnných . . . . .	20
1.4.3 Momentové odhady . . . . .	21
1.4.4 2SLS-odhad . . . . .	27
1.4.5 3SLS-odhad . . . . .	28
1.4.6 Maximálně věrohodné odhady . . . . .	30
1.4.7 Odhady $k$ -té třídy . . . . .	35
1.4.8 Shrnutí odhadových procedur . . . . .	36
1.5 Testy exogenity . . . . .	37
1.6 Testy specifikace . . . . .	38
1.7 Dynamické soustavy simultánních rovnic . . . . .	39
1.8 Předpovědi v rámci soustav simultánních rovnic . . . . .	42
1.8.1 Nedynamické modely . . . . .	42
1.8.2 Dynamické modely . . . . .	43
1.9 Výpočetní software . . . . .	45
<b>2 Základní pojmy životního pojištění</b>	<b>46</b>
<b>3 Ekonometrický model životní pojišťovny</b>	<b>50</b>
3.1 Formulace modelu . . . . .	50
3.2 Odhad parametrů modelu . . . . .	54
3.3 Použití bootstrapu . . . . .	57
3.4 Analýza scénářů . . . . .	62
<b>Závěr</b>	<b>66</b>
<b>A Bootstrapová procedura</b>	<b>67</b>
<b>Literatura</b>	<b>70</b>

Název práce: Ekonometrické soustavy simultánních rovnic v životním pojištění

Autor: Radek Hendrych

Katedra (ústav): Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: prof. RNDr. Tomáš Cipra, DrSc.

e-mail vedoucího: cipra@karlin.mff.cuni.cz

Abstrakt: V této práci se zabýváme teoretickými i praktickými otázkami souvisejícími s ekonometrickými soustavami (lineárních) simultánních rovnic. V první kapitole se seznámíme s teoretickými aspekty uvedené problematiky. Nemaľý prostor věnujeme odhadovým procedurám a srovnání jejich vlastností, neopomeneme zmínit ani otázky identifikace, nekonzistence OLS-odhadů při simultánním modelování, testování některých hypotéz specifických pro tuto oblast, dynamických soustav či konstrukce předpovědí v rámci příslušných modelů. Ve druhé kapitole představíme vybrané relevantní pojmy životního pojištění. Ve třetí kapitole ukážeme praktickou aplikaci teoretických poznatků na příkladu ekonometrického modelu finančních toků v životní pojišťovně operující na českém trhu. Porovnáme mezi sebou běžné odhadové postupy (2SLS a 3SLS přístup), provedeme některé testy, které nám poslouží k verifikaci vybraných informací o studovaném modelu. Uvedeme možnost využití residuálního bootstrapu, včetně ukázky použití při konstrukci intervalů spolehlivosti. Na závěr analyzujeme několik předpovědí v rámci odhadnutého modelu životní pojišťovny pro předem dané scénáře budoucího vývoje vybraných proměnných, což je z praktického hlediska velmi důležité.

Klíčová slova: ekonometrické modely, ekonometrické soustavy simultánních rovnic, model životní pojišťovny, residuální bootstrap.

Title: Econometric systems of simultaneous equations in life insurance

Author: Radek Hendrych

Department: Department of Probability and Mathematical Statistics

Supervisor: prof. RNDr. Tomáš Cipra, DrSc.

Supervisor's e-mail address: cipra@karlin.mff.cuni.cz

Abstract: In present work we deal with theoretical and practical issues related to econometric systems of (linear) simultaneous equations. In the first chapter we introduce to theoretical aspects of this problem. We devote considerable space to estimation procedures and comparisons of their properties, mention questions of identification, an inconsistency of OLS-estimates for the simultaneous modeling, tests of hypotheses specific to this area, dynamic systems and constructions of forecasts in models. In the second chapter we introduce selected basic concepts relevant to life insurance. In the third chapter we show the practical application of theoretical knowledge in the event of an econometric model of financial flows in the life insurance company operating on the Czech market. We compare ordinary estimation procedures (2SLS and 3SLS approach), perform some tests, which serve us to verify selected information on the studied model. We show the possibility of using residual bootstrap, including examples of use in the construction of confidence intervals. Finally we analyze several predictions of the estimated model of the life insurance company for predetermined scenarios for the development of selected variables, which is very important from practical point of view.

Keywords: econometric models, econometric systems of simultaneous equations, model of life insurance company, residual bootstrap.

# Úvod

V předložené práci se budeme podrobněji věnovat problematice ekonometrických soustav (lineárních) simultánních rovnic. Seznámíme se nejen s teoretickým zázemím a základními východisky tohoto konceptu, ale taktéž uvedeme příklad jeho aplikace v rámci modelu životní pojišťovny.

První kapitola textu představuje teoretický základ celé práce. Kromě samozřejmé exaktní formulace pojmu soustav (lineárních) simultánních rovnic včetně příslušných předpokladů obsahuje například pasáže řešící problematiku identifikace, nekonzistenci odhadu parametrů soustavy metodou nejmenších čtverců, otázky z oblasti testování některých vybraných hypotéz, dále upozorňuje na některá specifika dynamických soustav simultánních rovnic či na konstrukci předpovědí v rámci prezentovaných ekonometrických modelů. Patříčný důraz je kladen na přehled odhadových postupů a shrnutí jejich vlastností.

Ve druhé kapitole práce uvedeme základní pojmy z oblasti životního pojištění s ohledem na jejich využití při následné praktické aplikaci teorie ekonometrických soustav simultánních rovnic.

Třetí kapitola textu je, jak již bylo zmíněno výše, zaměřena na praktickou implementaci pojmů a postupů představených v předchozích částech diplomové práce. Konkrétně bude naším cílem analyzovat ekonometrický model finančních toků v životní pojišťovně operující na českém pojistném trhu. Například porovnáme výstupy běžně aplikovaných odhadových procedur, provedeme některé testy a ukážeme možnost využití statistického bootstrapu. Naznačíme rovněž jakým způsobem lze v intencích uvažovaného modelu predikovat hodnoty jednotlivých vysvětlovaných proměnných při daných scénářích budoucího vývoje.

Tématika ekonometrických soustav simultánních rovnic je obsažena v poměrně širokém spektru dostupné literatury, za všechny můžeme jmenovat česky psanou publikaci Cípra (2008) či cizojazyčné knihy Greene (2003) a Dhrymes (1994). Aplikacemi soustav simultánních rovnic v oblasti modelů životních pojišťoven se zabývají například články Cípra (1998) nebo Baranoff et al. (2007).

Na závěr připojme několik poznámek technického charakteru. Konvence zápisu a používání matematických symbolů je ve shodě s knihou Cípra (2008), případné výjimky či doplňky jsou v textu řádně komentovány. Konce důkazů označujeme symbolem  $\square$ , konce poznámek  $\diamond$  a konce příkladů pak  $\triangle$ .

# Kapitola 1

## Ekonometrické soustavy simultánních rovnic

V této kapitole se seznámíme s ekonometrickými soustavami (lineárních) simultánních rovnic, tyto se ve finanční a ekonometrické praxi objevují poměrně často, náleží totiž k obecnějšímu konceptu tzv. *víceroznicových ekonometrických soustav*, které současně analyzují větší počet vysvětlovaných proměnných. Poznamenejme, že tato úvodní pasáž je zpracována na základě Cipra (2008, kapitola 7).

Jednoduchým příkladem víceroznicového systému je dvourovnicová soustava, která prezentuje oboustranně kauzální vztah mezi proměnnými  $x$  a  $y$ , tudíž jedna z rovnic vysvětluje proměnnou  $y$  pomocí  $x$  (eventuálně za přítomnosti dalších vysvětlujících faktorů), druhá rovnice vysvětluje proměnnou  $x$  pomocí  $y$  (případně opět za přítomnosti dalších vysvětlujících faktorů).

Analýza víceroznicových soustav je rovněž typická z pohledu datových souborů, ty jsou v tomto případě tvořeny většími počty proměnných, které jsou v určitém časovém období sledovány jako časová řada (např. roční zisky všech pojišťoven na daném trhu, roční HDP vybraných států). Dochází tedy ke kombinaci průřezové (pro jednotlivé pojišťovny, pro jednotlivé státy) a časové složky (pro jednotlivá časová období). Data tohoto typu se často označují jako *poolová (pooled data)*.

Víceroznicové soustavy pak vyhovují následující formální charakterizaci. Příslušný datový soubor je tvořen časovými řadami (tj. posloupnostmi pozorování chronologicky řazených v čase) o délce  $T$  ( $t = 1, \dots, T$ ) pro  $m$  průřezových jednotek či veličin ( $j = 1, \dots, m$ ). Každá z vysvětlovaných proměnných  $y_{jt}$  (např. roční zisky pojišťoven v zemi) má být vysvětlena pomocí vektoru vysvětlujících proměnných  $\mathbf{x}_{jt}$ , jedná se o řádkový vektor s  $k-1$  složkami, intercept uvažujeme zvlášť. *Soustava (lineárních) ekonometrických rovnic* má potom obecně tvar

$$y_{jt} = \alpha_{jt} + \mathbf{x}_{jt} \boldsymbol{\gamma}_{jt} + \varepsilon_{jt}, \quad j = 1, \dots, m, \quad t = 1, \dots, T, \quad \text{var}(\boldsymbol{\varepsilon}) = \mathbf{E}(\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}') = \boldsymbol{\Sigma}, \quad (1.0.1)$$

kde  $\boldsymbol{\Sigma}$  je rozptylová matice ( $mT \times mT$ ) vektoru residuálních složek  $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}'_1, \dots, \boldsymbol{\varepsilon}'_m)'$ . Speciálně pro  $m = 1$  se vztah (1.0.1) za určitých podmínek redukuje na (zobecněný) model lineární regrese, ve kterém pokládáme  $\alpha_{1t} = \beta_1$  a  $\boldsymbol{\gamma}_{1t} = (\beta_2, \dots, \beta_k)'$ , podrobnosti nalezneme v Cipra (2008, kapitoly 3 a 4).

Model (1.0.1) je ovšem pro praktické použití příliš obecný, neboť obsahuje více parametrů, konkrétně  $kmT + mT(mT + 1)/2$ , než je počet pozorování, tj.  $mT$ .

V praxi se tak setkáváme se speciálními případy tohoto obecného vyjádření, a to zejména se *SUR soustavami, panelovými daty a soustavami simultánních rovnic*, bližší informace uvádí Cipra (2008, kapitola 7).

Na závěr této úvodní části se podrobněji zabýváme pojmy *exogenní* a *endogenní proměnné*, které jsou v rámci studovaných víceroznicových soustav značně důležité.

Obecně proměnnou  $x$  vysvětlující proměnnou  $y$  definujeme jako *exogenní*, jestliže se podmíněné rozdělení  $y$  za podmínky  $x$  nemění při změnách procesu generujícího  $x$ , tzv. *Data Generating Process (DGP)* je ekonometrický koncept zachycující skutečné vztahy v celé populaci pomocí obecného modelu s náhodnou složkou (viz např. Cipra (2008, odstavec 3.2)).

Například v rámci klasického lineárního regresního modelu

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim (0, \sigma^2 \mathbf{I}),$$

kde  $\mathbf{X}$  je matice ( $T \times k$ ),  $\boldsymbol{\beta}$  je vektor ( $k \times 1$ ) a  $\boldsymbol{\varepsilon}$  je vektor residuí ( $T \times 1$ ) (viz Cipra (2008, kapitola 3)), se definice exogenity transformuje na *podmínku současné nekorelovanosti regresorů s residuální složkou*, někdy také hovoříme o *podmínce ortogonality*, konkrétně  $\mathbf{E}(\mathbf{X}'\boldsymbol{\varepsilon}) = \mathbf{0}$  s asymptotickou verzí  $\text{plim}(\mathbf{X}'\boldsymbol{\varepsilon}/T) = \mathbf{0}$ , tímto vyjádřením rozumíme, že argument  $(\mathbf{X}'\boldsymbol{\varepsilon}/T)$  pro  $T \rightarrow \infty$  konverguje v pravděpodobnosti k  $\mathbf{0}$  (více uvádí Anděl (2007, dodatek B)).

Dále můžeme rozlišit, zda je exogenní proměnná

- *predeterminovaná*, tj. nekorelovaná v daném čase se současnými a budoucími hodnotami residuální složky, protože byla vytvořena danou soustavou v minulém čase;
- *striktně exogenní*, tj. nekorelovaná v daném čase se všemi hodnotami residuální složky, protože vznikla plně vně dané soustavy.

V zásadě je možné vycházet z představy, že exogenní proměnné jsou veličiny, které do dané soustavy vstupují zvnějšku nebo byly v jejím rámci vytvořeny v minulém čase, a mají tak z hlediska modelu v uvažovaném čase *externí charakter*.

Proměnné, které nesplňují definici exogenity, označujeme jako *endogenní*, neboť v dané soustavě v uvažovaném čase vznikají, a jsou tak jejím výstupem.

Uvažujme nyní jednoduchou dvourovnicovou soustavu

$$y_t = \alpha_1 + \beta_1 x_t + \varepsilon_{1t}, \tag{1.0.2}$$

$$x_t = \alpha_2 + \beta_2 y_t + \varepsilon_{2t}, \quad t = 1, \dots, T. \tag{1.0.3}$$

Vysvětlující proměnná  $x_t$  v první rovnici je endogenní, neboť vzhledem k přítomnosti druhé rovnice (1.0.3) závisí na proměnné  $y_t$ , potažmo také na residuální složce  $\varepsilon_{1t}$ . Z hlediska celé soustavy je pak proměnná  $x_t$  zjevně jejím výstupem. Tato tvrzení platí dokonce i v případě vzájemné nekorelovanosti residuálních složek  $\varepsilon_{1t}$  a  $\varepsilon_{2t}$ .

Jestliže by však proměnná  $x_t$  byla vytvořena soustavou v minulém čase, např. bychom v čase  $t$  v rovnici (1.0.2) uvažovali vysvětlující proměnnou typu  $x_{t-1}$  namísto proměnné  $x_t$ , potom by takový regresor za určitých podmínek kladených na residuální složky splňoval v rovnici (1.0.2) podmínky ortogonality, a byl tedy exogenní (ve smyslu predeterminovanosti).



Na základě předešlé diskuze je zřejmé, že endogenní proměnná musí aspoň v jedné rovnici dané soustavy vystupovat jako vysvětlovaná, zatímco exogenní proměnná může být vždy pouze vysvětlující (nesmí tedy v žádné rovnici příslušné soustavy působit jako vysvětlovaná).

Ve zbývajících částech kapitoly se budeme zabývat výhradně ekonometrickými soustavami lineárních simultánních rovnic, doplňující informace nalezneme například v publikacích Cipra (2008, kapitola 7) či Greene (2003, kapitola 15), podrobnějším zdrojem může být Dhrymes (1994). Problematika nelineárních modelů je velice dobře zachycena v Dhrymes (1994) nebo Davidson & MacKinnon (2004, kapitola 12).

## 1.1 Formulace soustavy simultánních rovnic

*Soustavy simultánních rovnic* (*simultaneous equation models, SEM*) patří historicky mezi první vícerovnicové ekonometrické modely, původně se používaly výhradně k modelování makroekonomických dat, např. HDP, spotřeby či vývozu. Velmi rozšířené bylo rovněž modelování celých národních ekonomik, například Tinbergův model nizozemské ekonomiky (1937) či Kleinův model hospodářství USA (1950).

Soustava simultánních rovnic je tvaru

$$y_{jt} = \sum_{i=1, i \neq j}^m \gamma_{ji} y_{it} + \sum_{i=1}^k \beta_{ji} x_{it} + \varepsilon_{jt}, \quad j = 1, \dots, m, \quad t = 1, \dots, T. \quad (1.1.1)$$

**Poznámka 1.1.1.** Uvedme několik dalších poznámek k soustavě (1.1.1).

1. Soustava obsahuje rovnici pro každou z  $m$  endogenních proměnných  $y_{jt}$ , dále obsahuje  $k$  exogenních proměnných  $x_{it}$ , včetně interceptu, který zde není formálně odlišen.
2. Pokud by platilo  $\gamma_{ji} = 0$  pro všechna  $i, j$ , pak by se soustava (1.1.1) zredukovala na SUR soustavu, podrobněji viz Cipra (2008, odstavec 7.2).
3. Soustavu (1.1.1) nazýváme simultánní, jestliže  $\gamma_{ij} \neq 0$  pro nějaké  $i \neq j$ . Jestliže pro nějaké  $i \neq j$  platí jak  $\gamma_{ji} \neq 0$ , tak  $\gamma_{ij} \neq 0$ , potom na sobě endogenní proměnné  $y_{jt}$  a  $y_{it}$  závisují oboustranně.

◇

Soustavu simultánních rovnic (1.1.1) lze zapsat maticově pro jednotlivé rovnice zvlášť jako

$$\mathbf{y}_j = \mathbf{Y}_j \boldsymbol{\gamma}_j + \mathbf{X}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j = \mathbf{Z}_j \boldsymbol{\delta}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, m, \quad (1.1.2)$$

kde  $\mathbf{y}_j$  je vektor ( $T \times 1$ ) pozorovaných hodnot  $j$ -té endogenní proměnné,  $\mathbf{Y}_j$  je matice ( $T \times (m_j - 1)$ ) pozorovaných hodnot endogenních proměnných na pravé straně  $j$ -té rovnice soustavy (1.1.1) s odpovídajícím vektorem  $((m_j - 1) \times 1)$  parametrů  $\boldsymbol{\gamma}_j$ ,  $\mathbf{X}_j$  je matice ( $T \times k_j$ ) pozorovaných hodnot exogenních proměnných  $j$ -té rovnice s odpovídajícím vektorem  $(k_j \times 1)$  parametrů  $\boldsymbol{\beta}_j$ ,  $\mathbf{Z}_j = (\mathbf{Y}_j, \mathbf{X}_j)$  je matice

$(T \times (k_j + m_j - 1))$  pozorovaných hodnot všech proměnných na pravé straně  $j$ -té rovnice soustavy (1.1.1) s odpovídajícím vektorem  $((k_j + m_j - 1) \times 1)$  parametrů  $\boldsymbol{\delta}_j = (\boldsymbol{\gamma}'_j, \boldsymbol{\beta}'_j)'$  a  $\boldsymbol{\varepsilon}_j$  je vektor  $(T \times 1)$  residuálních složek. Proměnné  $j$ -té rovnice soustavy (1.1.1) s nulovými parametry při přechodu k zápisu (1.1.2) ignorujeme, počet zbylých endogenních proměnných v  $j$ -té rovnici značíme  $m_j$  a počet zbylých exogenních proměnných  $k_j$ . Proměnná  $y_j$  z  $m_j$  endogenních proměnných  $j$ -té rovnice soustavy se nachází na levé straně této rovnosti, tudíž na pravé straně opravdu zbývá  $m_j - 1$  endogenních proměnných.

Soustavu simultánních rovnic (1.1.1) lze rovněž vyjádřit jediným maticovým vztahem pro všech  $m$  rovnic jako

$$\mathbf{Y}\boldsymbol{\Gamma} + \mathbf{X}\mathbf{B} + \mathbf{E} = \mathbf{Z}\boldsymbol{\Delta} + \mathbf{E} = \mathbf{0}, \quad (1.1.3)$$

kde  $\mathbf{Y}$  je matice  $(T \times m)$  pozorovaných hodnot všech  $m$  endogenních proměnných soustavy s odpovídající maticí  $(m \times m)$  parametrů  $\boldsymbol{\Gamma}$ ,  $\mathbf{X}$  je matice  $(T \times k)$  pozorovaných hodnot všech  $k$  exogenních proměnných soustavy s odpovídající maticí  $(k \times m)$  parametrů  $\mathbf{B}$ ,  $\mathbf{Z} = (\mathbf{Y}, \mathbf{X})$  je matice  $(T \times (k + m))$  pozorovaných hodnot všech  $(k + m)$  proměnných soustavy s odpovídající maticí  $((k + m) \times m)$  parametrů  $\boldsymbol{\Delta} = (\boldsymbol{\Gamma}', \mathbf{B}')$  a  $\mathbf{E} = (\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_m)$  je matice  $(T \times m)$  residuálních složek soustavy. Uvedme, že v tomto případě proměnné s nulovými parametry v jednotlivých rovnicích soustavy neignorujeme, a tak například  $j$ -tý sloupec matice  $\boldsymbol{\Gamma}$  není shodný s vektorem  $\boldsymbol{\gamma}_j$  s  $(m_j - 1)$  složkami, ale obsahuje  $m$  složek s případnými nulovými hodnotami.

Velmi často má matice  $(m \times m)$   $\boldsymbol{\Gamma}$  na diagonále hodnoty  $-1$ , uplatňuje se zde tzv. *normalizační pravidlo*, které usnadňuje převod soustavy ve tvaru (1.1.3) na tvar (1.1.2).

Poznamenejme, že tvar soustavy rovnic (1.1.3) se zpravidla označuje jako *strukturální tvar* soustavy simultánních rovnic.

Nyní uvedme obecné předpoklady, které se v souvislosti se soustavami simultánních rovnic obvykle uvažují:

**(P1)** Residuální složky mají nulovou střední hodnotu, jsou současně korelované, tj.  $\text{cov}(\varepsilon_{it}, \varepsilon_{jt}) = \mathbf{E}(\varepsilon_{it}\varepsilon_{jt}) = \sigma_{ij}$  pro všechna  $i, j, t$ , ale nejsou časově korelované, tj.  $\text{cov}(\varepsilon_{is}, \varepsilon_{jt}) = \mathbf{E}(\varepsilon_{is}\varepsilon_{jt}) = 0$  pro všechna  $i, j, s \neq t$ .

Kompaktněji lze tento předpoklad formulovat jako  $\text{var}(\boldsymbol{\varepsilon}'_t) = \mathbf{E}(\boldsymbol{\varepsilon}'_t \boldsymbol{\varepsilon}_t) = \tilde{\boldsymbol{\Sigma}} = \{\sigma_{ij}\}_{i,j=1,\dots,m}$  a  $\text{cov}(\boldsymbol{\varepsilon}'_s, \boldsymbol{\varepsilon}'_t) = \mathbf{E}(\boldsymbol{\varepsilon}'_s \boldsymbol{\varepsilon}_t) = \mathbf{0}$  pro  $s \neq t$ , přičemž  $\boldsymbol{\varepsilon}_t$  označuje  $t$ -tý řádek matice  $\mathbf{E}$ ,  $t = 1, \dots, T$ .

**(P2)**  $\mathbf{E}(\mathbf{x}'_t \boldsymbol{\varepsilon}_t) = \mathbf{E}(\mathbf{x}'_t) \mathbf{E}(\boldsymbol{\varepsilon}_t) = \mathbf{0}$  pro všechna  $t$ , analogicky  $\mathbf{x}_t$  označuje  $t$ -tý řádek matice  $\mathbf{X}$ ,  $t = 1, \dots, T$ .

**(P3)**  $\mathbf{E}(\mathbf{x}'_t \mathbf{x}_t) = \mathbf{Q}$  pro všechna  $t$ ,  $\mathbf{Q}$  je konečná regulární matice  $(k \times k)$ .

Asymptotické verze právě prezentovaných předpokladů **(P1)** - **(P3)** jsou následující:

**(A1)**  $\text{plim}(1/T) \mathbf{E}' \mathbf{E} = \tilde{\boldsymbol{\Sigma}}$ ,

$$(A2) \text{plim}(1/T)\mathbf{X}'\mathbf{E} = \mathbf{0},$$

$$(A3) \text{plim}(1/T)\mathbf{X}'\mathbf{X} = \mathbf{Q}.$$

Pro úplnost dodejme, že rozptylová matice ( $mT \times mT$ ) vektoru  $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}'_1, \dots, \boldsymbol{\varepsilon}'_m)'$  je na základě uvedených předpokladů tvaru

$$\text{var}(\boldsymbol{\varepsilon}) = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \begin{pmatrix} \sigma_{11}\mathbf{I}_T & \dots & \sigma_{1m}\mathbf{I}_T \\ \sigma_{21}\mathbf{I}_T & \dots & \sigma_{2m}\mathbf{I}_T \\ \vdots & \vdots & \vdots \\ \sigma_{m1}\mathbf{I}_T & \dots & \sigma_{mm}\mathbf{I}_T \end{pmatrix} = \tilde{\boldsymbol{\Sigma}} \otimes \mathbf{I}_T = \boldsymbol{\Sigma} \quad (1.1.4)$$

a že rovněž platí

$$E(\mathbf{E}'\mathbf{E}) = E(\{\boldsymbol{\varepsilon}'_i\boldsymbol{\varepsilon}_j\}_{i,j=1,\dots,m}) = T \begin{pmatrix} \sigma_{11} & \dots & \sigma_{1m} \\ \sigma_{21} & \dots & \sigma_{2m} \\ \vdots & \vdots & \vdots \\ \sigma_{m1} & \dots & \sigma_{mm} \end{pmatrix} = T\tilde{\boldsymbol{\Sigma}}. \quad (1.1.5)$$

Symbolem  $\otimes$  rozumíme *Kroneckerův součin matic*, konkrétně pro matici  $\mathbf{A}$  typu  $(p \times q)$  a matici  $\mathbf{B}$  typu  $(r \times s)$  je  $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$  matice  $(pr \times qs)$  skládající se z  $pq$  bloků, přičemž  $c_{ij} = a_{ij}\mathbf{B}$ ,  $i = 1, \dots, p$ ,  $j = 1, \dots, q$ .

**Poznámka 1.1.2.** V některých situacích můžeme předpokládat, že jsou exogenní proměnné nenáhodné, tj.  $\mathbf{x}_1, \dots, \mathbf{x}_T$  jsou nenáhodné vektory. V tomto případě lze pak namísto předpokladů (P2) - (P3) a (A2) - (A3) uvažovat pouze

$$\lim_{T \rightarrow \infty} \frac{\mathbf{X}'\mathbf{X}}{T} = \mathbf{Q}. \quad (1.1.6)$$

V mnoha modelech se však mezi exogenní proměnné řadí zpožděné endogenní proměnné (viz odstavec 1.7), které mají zřejmý stochastický charakter, tudíž budeme  $\mathbf{X}$  nadále chápat jako stochastický proces.  $\diamond$

Za předpokladu regularity čtvercové matice  $\boldsymbol{\Gamma}$  (tzv. *completeness condition*) je možné strukturální tvar soustavy (1.1.3) násobit zprava inverzní maticí  $\boldsymbol{\Gamma}^{-1}$ . Obdržíme tak tzv. *redukovaný tvar* soustavy simultánních rovnic

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\Pi} + \mathbf{V}, \text{ kde } \boldsymbol{\Pi} = -\mathbf{B}\boldsymbol{\Gamma}^{-1} \text{ a } \mathbf{V} = -\mathbf{E}\boldsymbol{\Gamma}^{-1}. \quad (1.1.7)$$

Redukovaný tvar skutečně ukazuje, že endogenní proměnné můžeme považovat za výstupy dané soustavy simultánních rovnic, neboť pro  $j$ -tou rovnici v (1.1.7) dostáváme

$$\mathbf{y}_{\cdot j} = \mathbf{X}\boldsymbol{\pi}_{\cdot j} + \mathbf{v}_{\cdot j}, \quad (1.1.8)$$

$\mathbf{y}_{\cdot j}$ ,  $\boldsymbol{\pi}_{\cdot j}$  a  $\mathbf{v}_{\cdot j}$  jsou  $j$ -té sloupce matic  $\mathbf{Y}$ ,  $\boldsymbol{\Pi}$  a  $\mathbf{V}$  ve vztahu (1.1.7). Snadno nahlédneme, že endogenní proměnné  $\mathbf{y}_{\cdot j}$  jsou lineární funkcí exogenních proměnných a residuální složky.

Z předpokladů **(P1)** - **(P3)**, resp. **(A1)** - **(A3)** obdržíme v rámci redukováného tvaru soustavy (1.1.7) následující vztahy:

$$\mathbf{E}(\mathbf{v}'_t) = -(\mathbf{\Gamma}^{-1})'\mathbf{E}(\boldsymbol{\varepsilon}'_t) = \mathbf{0}, \quad \forall t, \quad (1.1.9)$$

$$\text{var}(\mathbf{v}'_t) = \mathbf{E}(\mathbf{v}'_t \mathbf{v}_t) = (\mathbf{\Gamma}^{-1})'\mathbf{E}(\boldsymbol{\varepsilon}'_t \boldsymbol{\varepsilon}_t) \mathbf{\Gamma}^{-1} = \tilde{\boldsymbol{\Omega}}, \quad \forall t, \quad (1.1.10)$$

$$\text{cov}(\mathbf{v}'_s, \mathbf{v}'_t) = \mathbf{E}(\mathbf{v}'_s \mathbf{v}_t) = (\mathbf{\Gamma}^{-1})'\mathbf{E}(\boldsymbol{\varepsilon}'_s \boldsymbol{\varepsilon}_t) \mathbf{\Gamma}^{-1} = \mathbf{0}, \quad \forall s \neq t, \quad (1.1.11)$$

$$\mathbf{E}(\mathbf{x}'_t \mathbf{v}_t) = -\mathbf{E}(\mathbf{x}'_t \boldsymbol{\varepsilon}_t) \mathbf{\Gamma}^{-1} = \mathbf{0}, \quad \forall t, \quad (1.1.12)$$

$$\text{plim}(1/T) \mathbf{V}' \mathbf{V} = (\mathbf{\Gamma}^{-1})' [\text{plim}(1/T) \mathbf{E}' \mathbf{E}] \mathbf{\Gamma}^{-1} = \tilde{\boldsymbol{\Omega}}, \quad (1.1.13)$$

$$\text{plim}(1/T) \mathbf{X}' \mathbf{V} = -[\text{plim}(1/T) \mathbf{X}' \mathbf{E}] \mathbf{\Gamma}^{-1} = \mathbf{0}, \quad (1.1.14)$$

připomněme, že  $\mathbf{x}_t$ ,  $\mathbf{v}_t$  a  $\boldsymbol{\varepsilon}_t$  označují  $t$ -té řádky v maticích  $\mathbf{X}$ ,  $\mathbf{V}$  a  $\mathbf{E}$ .

Rozptylová matice  $(mT \times mT)$  vektoru  $\mathbf{v} = (\mathbf{v}'_1, \dots, \mathbf{v}'_m)'$ ,  $\mathbf{v}_j$  je  $j$ -tý sloupec matice  $\mathbf{V}$ , má na základě výše uvedených vlastností tvar

$$\text{var}(\mathbf{v}) = \mathbf{E}(\mathbf{v} \mathbf{v}') = (\mathbf{\Gamma}^{-1} \otimes \mathbf{I}_T)' \text{var}(\boldsymbol{\varepsilon}) (\mathbf{\Gamma}^{-1} \otimes \mathbf{I}_T) = (\mathbf{\Gamma}^{-1})' \tilde{\boldsymbol{\Sigma}} \mathbf{\Gamma}^{-1} \otimes \mathbf{I}_T = \tilde{\boldsymbol{\Omega}} \otimes \mathbf{I}_T, \quad (1.1.15)$$

neboť pro Kroneckerův součin matic platí

$$\begin{aligned} (\mathbf{A} \otimes \mathbf{B})' &= \mathbf{A}' \otimes \mathbf{B}', \\ (\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) &= (\mathbf{AC}) \otimes (\mathbf{BD}), \\ (\mathbf{A} \otimes \mathbf{B})^{-1} &= \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}, \text{ existují-li inverze.} \end{aligned}$$

Analogicky jako v (1.1.5) získáváme

$$\mathbf{E}(\mathbf{V}' \mathbf{V}) = \mathbf{E}(\{\mathbf{v}'_i \mathbf{v}_j\}_{i,j=1,\dots,m}) = T \tilde{\boldsymbol{\Omega}}. \quad (1.1.16)$$

## 1.2 Problematika identifikace

*Problém identifikace* spočívá ve specifikaci podmínek, za kterých jsme schopni nalézt (jednoznačný) vztah mezi parametry redukováného tvaru soustavy (1.1.7), tj.  $(\boldsymbol{\Pi}, \tilde{\boldsymbol{\Omega}})$ , a parametry strukturálního tvaru soustavy (1.1.3), tj.  $(\boldsymbol{\Gamma}, \mathbf{B}, \tilde{\boldsymbol{\Sigma}})$ .

Mějme strukturální tvar soustavy (1.1.3)

$$\mathbf{Y} \boldsymbol{\Gamma} + \mathbf{X} \mathbf{B} + \mathbf{E} = \mathbf{0}, \quad (1.2.1)$$

tento vynásobíme zprava libovolnou nenulovou regulární čtvercovou maticí  $\mathbf{F}$  typu  $(m \times m)$ , dostaneme

$$\mathbf{Y} \bar{\boldsymbol{\Gamma}} + \mathbf{X} \bar{\mathbf{B}} + \bar{\mathbf{E}} = \mathbf{0}, \quad (1.2.2)$$

kde  $\bar{\boldsymbol{\Gamma}} = \boldsymbol{\Gamma} \mathbf{F}$ ,  $\bar{\mathbf{B}} = \mathbf{B} \mathbf{F}$  a  $\bar{\mathbf{E}} = \mathbf{E} \mathbf{F}$ , přičemž  $\text{var}(\bar{\boldsymbol{\varepsilon}}) = \mathbf{F}' \tilde{\boldsymbol{\Sigma}} \mathbf{F} \otimes \mathbf{I}_T$  (analogie (1.1.4)).

Snadno nahlédneme, že redukováné tvary soustav (1.2.1) a (1.2.2) jsou ekvivalentní (*observationally equivalent*), neboť platí

$$-\mathbf{B} \boldsymbol{\Gamma}^{-1} = \boldsymbol{\Pi} = -\bar{\mathbf{B}} \bar{\boldsymbol{\Gamma}}^{-1} = -\mathbf{B} \mathbf{F} \mathbf{F}^{-1} \boldsymbol{\Gamma}^{-1}, \quad -\mathbf{E} \boldsymbol{\Gamma}^{-1} = \mathbf{V} = -\bar{\mathbf{E}} \bar{\boldsymbol{\Gamma}}^{-1} = -\mathbf{E} \mathbf{F} \mathbf{F}^{-1} \boldsymbol{\Gamma}^{-1},$$

$$(\boldsymbol{\Gamma}^{-1})' \tilde{\boldsymbol{\Sigma}} \boldsymbol{\Gamma}^{-1} \otimes \mathbf{I}_T = \text{var}(\mathbf{v}) = (\bar{\boldsymbol{\Gamma}}^{-1})' \text{var}(\bar{\boldsymbol{\varepsilon}}) \bar{\boldsymbol{\Gamma}}^{-1} \otimes \mathbf{I}_T.$$

Odtud jasně vyplývá, že bez určitých omezení kladených na parametry soustavy není ve smyslu předchozí diskuze možné nalézt jednoznačný vztah mezi trojicí strukturálních parametrů  $(\mathbf{\Gamma}, \mathbf{B}, \tilde{\mathbf{\Sigma}})$  a dvojicí parametrů redukovaného tvaru  $(\mathbf{\Pi}, \tilde{\mathbf{\Omega}})$ .

Přípustnou transformací  $\mathbf{F}$  pro trojici strukturálních parametrů  $(\mathbf{\Gamma}, \mathbf{B}, \tilde{\mathbf{\Sigma}})$  rozumíme takovou nenulovou regulární matici  $\mathbf{F}$  typu  $(m \times m)$ , že  $\Delta\mathbf{F} = (\mathbf{\Gamma}', \mathbf{B}')'\mathbf{F}$  vyhovuje všem (známým) restrikcím kladeným na  $\Delta$  a  $\mathbf{F}'\tilde{\mathbf{\Sigma}}\mathbf{F}$  splňuje všechna (známá) omezení kladená na  $\tilde{\mathbf{\Sigma}}$ .

Řekneme, že  $j$ -tá rovnice strukturálního tvaru soustavy (1.1.3) je *identifikovaná*, jestliže všechny její ekvivalentní (*observationally equivalent*) redukované tvary vznikly pomocí přípustných transformací  $\mathbf{F}$  takových, že jejich  $j$ -té sloupce jsou tvořeny vektory tvaru  $\alpha_j \mathbf{e}_j$ ,  $\alpha_j \neq 0$ ;  $\mathbf{e}_1 = (1, 0, 0, \dots)'$ ,  $\mathbf{e}_2 = (0, 1, 0, \dots)'$ , atp. Řekneme, že soustava ve strukturálním tvaru (1.1.3) je *identifikovaná*, jestliže je každá její rovnice identifikovaná. V případě, že je v  $j$ -té rovnici uplatněno normalizační pravidlo, platí  $\alpha_j = 1$ ,  $j = 1, \dots, m$ .

Problematika identifikace je nejčastěji řešena využitím apriorní znalosti o přítomnosti či nepřítomnosti parametrů v jednotlivých rovnicích strukturálního tvaru soustavy.

Vezměme nyní v úvahu  $j$ -tou rovnici strukturálního tvaru soustavy (1.1.3)

$$\mathbf{Y}\boldsymbol{\gamma}_{\cdot j} + \mathbf{X}\boldsymbol{\beta}_{\cdot j} + \boldsymbol{\varepsilon}_{\cdot j} = \mathbf{0}, \quad j = 1, \dots, m, \quad (1.2.3)$$

kde  $\boldsymbol{\gamma}_{\cdot j}$ ,  $\boldsymbol{\beta}_{\cdot j}$  a  $\boldsymbol{\varepsilon}_{\cdot j}$  označují  $j$ -té sloupce příslušných matic  $\mathbf{\Gamma}$ ,  $\mathbf{B}$  a  $\mathbf{E}$ .

Předpokládejme, že vektory  $\boldsymbol{\gamma}_{\cdot j}$  a  $\boldsymbol{\beta}_{\cdot j}$  jsou při zohlednění apriorní informace o nepřítomnosti některých parametrů (po případném přerovnání) ve tvaru

$$\boldsymbol{\gamma}_{\cdot j} = \begin{pmatrix} \boldsymbol{\gamma}_{\cdot j}^* \\ \mathbf{0} \end{pmatrix}, \quad \boldsymbol{\beta}_{\cdot j} = \begin{pmatrix} \boldsymbol{\beta}_{\cdot j}^* \\ \mathbf{0} \end{pmatrix},$$

přičemž vektor  $\boldsymbol{\gamma}_{\cdot j}^*$  má  $m_j$  složek a vektor  $\boldsymbol{\beta}_{\cdot j}^*$  má  $k_j$  složek.

Následně můžeme příslušným způsobem rozdělit rovněž matice  $\mathbf{\Gamma}$  a  $\mathbf{B}$ , tj.

$$\mathbf{\Gamma} = \begin{pmatrix} \boldsymbol{\gamma}_{\cdot j}^* & \mathbf{\Gamma}_{12} \\ \mathbf{0} & \mathbf{\Gamma}_{22}^j \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \boldsymbol{\beta}_{\cdot j}^* & \mathbf{B}_{12} \\ \mathbf{0} & \mathbf{B}_{22}^j \end{pmatrix}$$

Uvažujme nyní přípustnou transformaci  $\mathbf{F}$ . Platí

$$\Delta\mathbf{F} = \begin{pmatrix} \mathbf{\Gamma} \\ \mathbf{B} \end{pmatrix} \mathbf{F} = \begin{pmatrix} \boldsymbol{\gamma}_{\cdot j}^* & \mathbf{\Gamma}_{12} \\ \mathbf{0} & \mathbf{\Gamma}_{22}^j \\ \boldsymbol{\beta}_{\cdot j}^* & \mathbf{B}_{12} \\ \mathbf{0} & \mathbf{B}_{22}^j \end{pmatrix} \cdot \begin{pmatrix} f_{11} & \mathbf{f}'_{12} \\ \mathbf{f}_{21} & \mathbf{F}_{22} \end{pmatrix},$$

kde  $f_{11} \in \mathbb{R}$ ,  $\mathbf{f}_{12}$ ,  $\mathbf{f}_{21}$  jsou vektory s  $(m - 1)$  složkami a  $\mathbf{F}_{22}$  je čtvercová matice o rozměrech  $((m - 1) \times (m - 1))$ .

Přípustná transformace  $\mathbf{F}$  musí zohledňovat veškerá omezení, která jsou kladena na strukturální parametry  $\Delta$ , tudíž je nutně splněna rovnost

$$\begin{pmatrix} \mathbf{\Gamma}_{22}^j \\ \mathbf{B}_{22}^j \end{pmatrix} \mathbf{f}_{21} = \mathbf{0}. \quad (1.2.4)$$

Víme, že  $j$ -tá rovnice soustavy je identifikovaná, právě když platí

$$\mathbf{f}_{21} = \mathbf{0}. \quad (1.2.5)$$

Předpokládejme, že

$$h \left( \begin{array}{c} \mathbf{\Gamma}_{22}^j \\ \mathbf{B}_{22}^j \end{array} \right) = m - 1. \quad (1.2.6)$$

Potom ze vztahu (1.2.4) vyplývá rovnost (1.2.5), neboť matice v rovnici (1.2.4) má  $m - 1$  sloupců, a tak je v tomto případě  $j$ -tá rovnice soustavy identifikovaná.

Nyní naopak uvažujme, že je  $j$ -tá rovnice soustavy identifikovaná a současně

$$h \left( \begin{array}{c} \mathbf{\Gamma}_{22}^j \\ \mathbf{B}_{22}^j \end{array} \right) < m - 1,$$

potom existuje nenulový vektor  $\tilde{\mathbf{f}}_{21}$  vyhovující rovnosti (1.2.4). Dále pro  $\tilde{f}_{11} \neq 0$  vezměme matici

$$\tilde{\mathbf{F}} = \left( \begin{array}{cc} \tilde{f}_{11} & \mathbf{0}' \\ \tilde{\mathbf{f}}_{21} & \mathbf{I} \end{array} \right).$$

K tomu, aby matice  $\tilde{\mathbf{F}}$  byla přípustnou transformací, musí být regulární, což je snadno viditelné, a musí zohlednit veškerá uvažovaná omezení kladená na strukturální parametry  $\mathbf{\Delta}$ , v našem případě se pak opět jedná o splnění vztahu (1.2.4) pro vektor  $\tilde{\mathbf{f}}_{21}$ . Tato situace ovšem vede ke sporu s identifikací dané rovnice, neboť první sloupec přípustné transformace  $\tilde{\mathbf{F}}$  je různý od vektoru  $\alpha \mathbf{e}_1$ ,  $\alpha \neq 0$ .

Je tedy zřejmé, že nutnou a postačující podmínkou identifikace  $j$ -té rovnice soustavy je vztah (1.2.6), tuto nazýváme *hodnostní podmínkou identifikace (rank condition)*.

Vzhledem k tomu, že matice  $((\mathbf{\Gamma}_{22}^j)', (\mathbf{B}_{22}^j)')$  má  $(m - m_j) + (k - k_j)$  řádků, potom s ohledem na hodnostní podmínku (1.2.6) nutně platí

$$(m - m_j) + (k - k_j) \geq m - 1, \quad (1.2.7)$$

tuto nutnou podmínku identifikace  $j$ -té rovnice soustavy nazýváme *rozměrovou podmínkou identifikace (order condition)*.

Na základě předchozí diskuze lze  $j$ -tou rovnici soustavy simultánních rovnic zařadit právě do jedné ze tří skupin:

1. *neidentifikovaná (unidentified)*

Tato situace nastává, jestliže rovnice nevyhovuje hodnostní podmínce nebo

$$(m - m_j) + (k - k_j) < m - 1, \quad (1.2.8)$$

prezentovanou nerovnost můžeme interpretovat tak, že z  $j$ -té strukturální rovnice soustavy nebyl odstraněn dostatek proměnných vzhledem k jejich celkovému počtu.

2. *přesně identifikovaná (exactly identified)*

Tato situace nastává, jestliže rovnice vyhovuje hodnostní podmínce a současně

$$(m - m_j) + (k - k_j) = m - 1. \quad (1.2.9)$$

### 3. přeidentifikovaná (*overidentified*)

Tato situace nastává, jestliže rovnice vyhovuje hodnotní podmínce a současně

$$(m - m_j) + (k - k_j) > m - 1, \quad (1.2.10)$$

uvedenou nerovnost je možné chápat tak, že z  $j$ -té strukturální rovnice soustavy bylo odstraněno nadbytečné množství proměnných.

Vraťme se opět k redukovanému tvaru soustavy. Ze vztahu  $\mathbf{\Pi} = -\mathbf{B}\mathbf{\Gamma}^{-1}$  s ohledem na předešlé značení vyplývá

$$\begin{pmatrix} \mathbf{\Pi}_{11}^j & \mathbf{\Pi}_{12} \\ \mathbf{\Pi}_{21}^j & \mathbf{\Pi}_{22} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\gamma}_{\cdot j}^* & \mathbf{\Gamma}_{12} \\ \mathbf{0} & \mathbf{\Gamma}_{22}^j \end{pmatrix} = - \begin{pmatrix} \boldsymbol{\beta}_{\cdot j}^* & \mathbf{B}_{12} \\ \mathbf{0} & \mathbf{B}_{22}^j \end{pmatrix}, \quad (1.2.11)$$

matice  $(k \times m)$  parametrů  $\mathbf{\Pi}$  je složená z matic  $\mathbf{\Pi}_{11}^j$  o rozměrech  $(k_j \times m_j)$ ,  $\mathbf{\Pi}_{12}$  o rozměrech  $(k_j \times (m - m_j))$ ,  $\mathbf{\Pi}_{21}^j$  o rozměrech  $((k - k_j) \times m_j)$  a  $\mathbf{\Pi}_{22}$  o rozměrech  $((k - k_j) \times (m - m_j))$ .

Odtud snadno nahlédneme, že

$$-\mathbf{\Pi}_{11}^j \boldsymbol{\gamma}_{\cdot j}^* = \boldsymbol{\beta}_{\cdot j}^*, \quad (1.2.12)$$

$$\mathbf{\Pi}_{21}^j \boldsymbol{\gamma}_{\cdot j}^* = \mathbf{0}. \quad (1.2.13)$$

Řešení soustavy rovnic (1.2.12) je plně determinováno řešením soustavy (1.2.13), která je jednoznačně řešitelná (až na násobek skalárem), právě když  $h(\mathbf{\Pi}_{21}^j) = m_j - 1$ , naprostou jednoznačnost nám poté poskytne normalizační pravidlo. Dhrymes (1994, odstavec 3.2.5) ukázal, že uvedená podmínka je ekvivalentní s hodnotní podmínkou (1.2.6).

Doplňme, že v případě  $h(\mathbf{\Pi}_{21}^j) < m_j - 1$  nejsme schopni (ani s využitím normalizačního pravidla) nalézt jednoznačně určené řešení soustavy (1.2.13).

Poslední možná situace, tj.  $h(\mathbf{\Pi}_{21}^j) = m_j$ , nás podle Dhrymes (1994, odstavec 4.8) vede k závěru, že redukovaný tvar soustavy simultánních rovnic není pomocí známých omezení vhodně zformulován.

**Příklad 1.2.1.** Uvažujme následující třírovnícovou soustavu simultánních rovnic

$$y_{1t} = \gamma_{12}y_{2t} + \gamma_{13}y_{3t} + \beta_{11} + \beta_{13}x_{3t} + \varepsilon_{1t}, \quad (1.2.14)$$

$$y_{2t} = \gamma_{21}y_{1t} + \beta_{21} + \varepsilon_{2t}, \quad (1.2.15)$$

$$y_{3t} = \beta_{31} + \beta_{32}x_{2t} + \beta_{33}x_{3t} + \beta_{34}x_{4t} + \varepsilon_{3t}. \quad (1.2.16)$$

Zřejmě platí  $m = 3$ ,  $k = 4$ . Snadno rovněž zjistíme tvar matic parametrů  $\mathbf{\Gamma}$  a  $\mathbf{B}$ , tj.

$$\mathbf{\Gamma} = \begin{pmatrix} -1 & \gamma_{21} & 0 \\ \gamma_{12} & -1 & 0 \\ \gamma_{13} & 0 & -1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ 0 & 0 & \beta_{32} \\ \beta_{13} & 0 & \beta_{33} \\ 0 & 0 & \beta_{34} \end{pmatrix}.$$

V první rovnici (1.2.14) jednoduše spočteme

$$(m - m_1) + (k - k_1) = (3 - 3) + (4 - 2) = 2 = m - 1 = 2,$$

takže je splněna nutná podmínka identifikace. Po vhodném přerovnání dostáváme

$$\mathbf{\Gamma}_{22}^1 = \begin{pmatrix} 0 & 0 \end{pmatrix}, \quad \mathbf{B}_{22}^1 = \begin{pmatrix} 0 & \beta_{32} \\ 0 & \beta_{34} \end{pmatrix} \Rightarrow h \begin{pmatrix} \mathbf{\Gamma}_{22}^1 \\ \mathbf{B}_{22}^1 \end{pmatrix} < m - 1 = 2.$$

První rovnice soustavy (1.2.14) tedy v tomto případě nevyhovuje hodnotní podmínce identifikace. Tato rovnice je tudíž neidentifikovaná.

Ve druhé rovnici (1.2.15) vychází

$$(m - m_2) + (k - k_2) = (3 - 2) + (4 - 1) = 4 > m - 1 = 2,$$

takže vyhovuje rozměrové podmínce identifikace, po přerovnání získáváme

$$\mathbf{\Gamma}_{22}^2 = \begin{pmatrix} \gamma_{13} & -1 \end{pmatrix}, \quad \mathbf{B}_{22}^2 = \begin{pmatrix} 0 & \beta_{32} \\ \beta_{13} & \beta_{33} \\ 0 & \beta_{34} \end{pmatrix}.$$

Hodnotní podmínka identifikace je splněna, právě když  $\beta_{13} \neq 0$  a současně  $\beta_{32} \neq 0$  nebo  $\beta_{34} \neq 0$ , v tomto případě je druhá rovnice soustavy (1.2.15) předidentifikovaná.

Třetí rovnice (1.2.16) je přesně identifikovaná, neboť neobsahuje žádné endogenní vysvětlující proměnné a  $k_3 = 4$ .  $\triangle$

V předchozích částech tohoto odstavce jsme na problematiku identifikace nahlíželi na základě apriorní znalosti o přítomnosti či nepřítomnosti parametrů v jednotlivých rovnicích strukturálního tvaru soustavy (tzv. *exclusion restrictions*). Podívejme se nyní na příklad, ve kterém bude identifikace řešena na základě informace o varianční matici  $\tilde{\Sigma}$ .

Uvažujme *rekurzivní soustavu* (*recursive system, triangular system*)

$$\begin{aligned} y_{1t} &= \mathbf{x}_t \cdot \boldsymbol{\beta}_1 + \varepsilon_{1t}, \\ y_{2t} &= \gamma_{21} y_{1t} + \mathbf{x}_t \cdot \boldsymbol{\beta}_2 + \varepsilon_{2t}, \\ &\vdots \\ y_{mt} &= \gamma_{m1} y_{1t} + \gamma_{m2} y_{2t} + \cdots + \gamma_{m,m-1} y_{m-1,t} + \mathbf{x}_t \cdot \boldsymbol{\beta}_m + \varepsilon_{mt}, \end{aligned} \tag{1.2.17}$$

vidíme, že v každé rovnici jako endogenní proměnné vystupují vysvětlované proměnné z předcházejících rovnic. První rovnice soustavy neobsahuje žádnou endogenní vysvětlující proměnnou, je tudíž identifikovaná a navíc lze standardním způsobem konzistentně odhadnout metodou nejmenších čtverců (OLS-metodou). Ostatní rovnice soustavy však nejsou bez doplňujících omezení identifikované, jak ukážeme níže.

Zdůrazněme, že zapíšeme-li uvedenou rekurzivní soustavu (1.2.17) ve strukturálním tvaru (1.1.3), matice  $\mathbf{\Gamma}$  bude zjevně horní trojúhelníková s hodnotami  $-1$  na diagonále (vzhledem k uplatnění normalizačního pravidla).

Předpokládejme nyní, že matice  $\tilde{\Sigma}$  je diagonální, tj.  $\text{cov}(\varepsilon_{it}, \varepsilon_{jt}) = 0$  pro všechna  $t$  a  $i \neq j$ . Soustava rovnic (1.2.17) se spolu s tímto předpokladem označuje jako *plně rekurzivní soustava* (*fully recursive system*). Za uvedené premisy jsou v  $j$ -té rovnici proměnné na její pravé straně nekorelované s residuální složkou  $\varepsilon_{jt}$ , toto jednoduše



plyne z „trojúhelníkového“ tvaru soustavy. Každá rovnice této soustavy může být důsledkem tohoto konzistentně odhadnuta OLS-metodou (OLS-odhad parametrů strukturálního tvaru soustavy simultánních rovnic nemusí být obecně konzistentní, viz následující odstavec 1.3).

Vraťme se nyní k problematice identifikace rekurzivní soustavy. Vezměme nenulovou čtvercovou matici  $\mathbf{F} = (\mathbf{f}_{\cdot 1}, \dots, \mathbf{f}_{\cdot m})$  typu  $(m \times m)$ . K tomu, aby se jednalo o přípustnou transformaci, musí být  $\mathbf{F}$  regulární a zohlednit veškerá omezení kladená na trojici strukturálních parametrů  $(\mathbf{\Gamma}, \mathbf{B}, \tilde{\mathbf{\Sigma}})$ . Vzhledem ke tvaru matice  $\mathbf{\Gamma}$  (viz výše) a identifikaci první rovnice soustavy je zřejmé, že  $\mathbf{F}$  musí být horní trojúhelníková matice (splňuje podmínku regularity) a  $\mathbf{f}_{\cdot 1} = \mathbf{e}_1$ . Potvrzuje se tak, že bez dalších apriorních informací jsou všechny rovnice rekurzivní soustavy (s výjimkou první) neidentifikované.

Při formulaci plně rekurzivní soustavy jsme předpokládali, že matice  $\tilde{\mathbf{\Sigma}}$  je diagonální. Přípustná lineární transformace  $\mathbf{F}$  tak musí zachovávat nulové hodnoty v  $\mathbf{F}'\tilde{\mathbf{\Sigma}}\mathbf{F}$  ve shodě s  $\tilde{\mathbf{\Sigma}}$ .

Odtud plyne  $\mathbf{f}'_{\cdot 1}\tilde{\mathbf{\Sigma}}\mathbf{f}_{\cdot 2} = 0$ . Víme, že  $\mathbf{f}_{\cdot 1} = \mathbf{e}_1$ , a tak  $\mathbf{f}'_{\cdot 1}\tilde{\mathbf{\Sigma}}\mathbf{f}_{\cdot 2} = \sigma_{11}f_{12} = 0$ , tj.  $f_{12} = 0$ . Vzhledem k uplatnění normalizačního pravidla pak v druhé rovnici soustavy platí  $\mathbf{f}_{\cdot 2} = \mathbf{e}_2$ .

Podobně obdržíme rovnost pro  $\mathbf{f}_{\cdot 3}$ . Pro přípustnou transformaci  $\mathbf{F}$  nutně platí  $\mathbf{f}'_{\cdot 1}\tilde{\mathbf{\Sigma}}\mathbf{f}_{\cdot 3} = 0$  a současně  $\mathbf{f}'_{\cdot 2}\tilde{\mathbf{\Sigma}}\mathbf{f}_{\cdot 3} = 0$ , z čehož na základě analogické argumentace jako výše dostaneme  $\mathbf{f}_{\cdot 3} = \mathbf{e}_3$ .

Pokračujeme-li dále, získáme  $\mathbf{F} = \mathbf{I}_m$ , což znamená, že plně rekurzivní soustava je identifikovaná.

**Poznámka 1.2.1.** V literatuře, viz například Dhrymes (1994, odstavec 3.2), se vedle dvou právě prezentovaných přístupů k problematice identifikace setkáváme ještě s několika dalšími, např. je možné využít apriorní informace o vzájemných vztazích mezi parametry z různých rovnic soustavy či znalost obecnějších lineárních omezení parametrů soustavy.  $\diamond$

**Poznámka 1.2.2.** V ekonomických a finančních aplikacích se nejčastěji vyskytují přeidentifikované rovnice, tudíž zkoumání podmínek identifikace ztrácí na významu.  $\diamond$

### 1.3 Nekonzistence OLS-odhadu v rámci simultánního modelování

V tomto odstavci poukážeme na nevhodnost použití klasické metody nejmenších čtverců při odhadování parametrů v soustavách simultánních rovnic. OLS-přístup totiž v tomto případě není obecně konzistentní. Nekonzistence odhadu je způsobena porušením asymptotických podmínek ortogonalit ze strany endogenních vysvětlujících proměnných.

Jako ilustrace zmíněného problému nám poslouží dvourovnicová soustava simultánních rovnic

$$y_{1t} = \gamma_1 y_{2t} + \beta_1 + \varepsilon_{1t}, \quad (1.3.1)$$

$$y_{2t} = \gamma_2 y_{1t} + \beta_2 + \varepsilon_{2t}, \quad (1.3.2)$$

mezi endogenními proměnnými  $y_1$  a  $y_2$  existuje oboustranná závislost, v obou rovnicích je obsažena právě jedna striktně exogenní proměnná, a to intercept.

Podívejme se nyní podrobněji na první rovnici soustavy (1.3.1). Pro OLS-odhad  $c_1$  parametru  $\gamma_1$  odvodíme ze soustavy normálních rovnic (viz Cipra (2008, odstavec 3.2)) rovnost

$$c_1 \cdot \left( \sum_{t=1}^T (y_{2t} - \bar{y}_2)^2 \right) = \sum_{t=1}^T y_{1t} y_{2t} - T \bar{y}_1 \bar{y}_2,$$

odtud dosazením vztahu (1.3.1) a jednoduchými úpravami obdržíme

$$c_1 \cdot \left( \sum_{t=1}^T (y_{2t} - \bar{y}_2)^2 \right) = \gamma_1 \cdot \left( \sum_{t=1}^T (y_{2t} - \bar{y}_2)^2 \right) + \sum_{t=1}^T (y_{2t} - \bar{y}_2) \varepsilon_{1t},$$

kde  $\bar{y}_1 = \frac{1}{T} \sum_{t=1}^T y_{1t}$ ,  $\bar{y}_2 = \frac{1}{T} \sum_{t=1}^T y_{2t}$ .

Odhad  $c_1$  lze tudíž za podmínky, že proměnné  $y_{2t}$  pro všechna  $t$  nenabývají stejných hodnot, vyjádřit jako

$$c_1 = \gamma_1 + \frac{\sum_{t=1}^T (y_{2t} - \bar{y}_2)(\varepsilon_{1t} - \bar{\varepsilon}_1)}{\sum_{t=1}^T (y_{2t} - \bar{y}_2)^2}, \quad (1.3.3)$$

kde  $\bar{\varepsilon}_1 = \frac{1}{T} \sum_{t=1}^T \varepsilon_{1t}$ .

Využijeme-li předchozí vztah (1.3.3) pro odhad  $c_1$ , nahlédneme

$$\text{plim } c_1 = \gamma_1 + \text{plim } \frac{\sum_{t=1}^T (y_{2t} - \bar{y}_2)(\varepsilon_{1t} - \bar{\varepsilon}_1)}{\sum_{t=1}^T (y_{2t} - \bar{y}_2)^2} \approx \gamma_1 + \frac{\text{cov}(y_{2t}, \varepsilon_{1t})}{\text{var}(y_{2t})} \neq \gamma_1, \quad (1.3.4)$$

přičemž jsme aproximativně použili slabý zákon velkých čísel (viz Anděl (2007, dodatek B)). Nekonzistence odhadu  $c_1$  parametru  $\gamma_1$  je způsobena přítomností druhé rovnice soustavy (1.3.2), neboť platí

$$\text{plim } \frac{1}{T} \sum_{t=1}^T y_{2t} \varepsilon_{1t} \neq 0.$$

V této souvislosti hovoříme o *nekonzistenci OLS-odhadu v důsledku simultánního modelování*.

K odhadu parametrů soustav simultánních rovnic zjevně klasický OLS-přístup nestačí.

## 1.4 Odhady parametrů v rámci soustav simultánních rovnic

### 1.4.1 Nepřímý odhad metodou nejmenších čtverců

Uvažujme redukováný tvar soustavy simultánních rovnic (1.1.7)

$$\mathbf{Y} = \mathbf{X}\mathbf{\Pi} + \mathbf{V}, \text{ kde } \mathbf{\Pi} = -\mathbf{B}\mathbf{\Gamma}^{-1} \text{ a } \mathbf{V} = -\mathbf{E}\mathbf{\Gamma}^{-1}. \quad (1.4.1)$$

Z uvedených předpokladů (A1) - (A3) vyplývá

$$\begin{aligned} \text{plim}(1/T)\mathbf{X}'\mathbf{Y} &= \text{plim}(1/T)\mathbf{X}'(\mathbf{X}\mathbf{\Pi} + \mathbf{V}) = \mathbf{Q}\mathbf{\Pi}, \\ \text{plim}(1/T)\mathbf{Y}'\mathbf{Y} &= \text{plim}(1/T)(\mathbf{X}\mathbf{\Pi} + \mathbf{V})'(\mathbf{X}\mathbf{\Pi} + \mathbf{V}) = \mathbf{\Pi}'\mathbf{Q}\mathbf{\Pi} + \tilde{\mathbf{\Omega}}. \end{aligned}$$

Odtud snadno nahlédneme

$$\begin{aligned} \mathbf{\Pi} &= \left[ \text{plim} \left( \frac{\mathbf{X}'\mathbf{X}}{T} \right) \right]^{-1} \left[ \text{plim} \left( \frac{\mathbf{X}'\mathbf{Y}}{T} \right) \right], \\ \tilde{\mathbf{\Omega}} &= \text{plim} \left( \frac{\mathbf{Y}'\mathbf{Y}}{T} \right) - \text{plim} \left( \frac{\mathbf{Y}'\mathbf{X}}{T} \right) \left( \frac{\mathbf{X}'\mathbf{X}}{T} \right)^{-1} \left( \frac{\mathbf{X}'\mathbf{Y}}{T} \right). \end{aligned}$$

Je tedy zřejmé, že obě matice  $\mathbf{\Pi}$  a  $\tilde{\mathbf{\Omega}}$  mohou být konzistentně odhadnuty metodou nejmenších čtverců, konkrétně OLS-odhad  $\mathbf{P}$  parametru  $\mathbf{\Pi}$  je

$$\mathbf{P} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}. \quad (1.4.2)$$

Asymptotická rozptylová matice vektoru  $\mathbf{p} = (\mathbf{p}'_1, \dots, \mathbf{p}'_m)'$ , kde  $\mathbf{p}_j$  označuje  $j$ -tý sloupec matice  $\mathbf{P}$ , je tvaru

$$\frac{\tilde{\mathbf{\Omega}} \otimes \mathbf{Q}^{-1}}{T}, \quad (1.4.3)$$

jejím konzistentním odhadem je

$$\mathbf{S} \otimes (\mathbf{X}'\mathbf{X})^{-1}, \quad \mathbf{S} = \frac{(\mathbf{Y} - \mathbf{X}\mathbf{P})'(\mathbf{Y} - \mathbf{X}\mathbf{P})}{T}. \quad (1.4.4)$$

Matici  $\mathbf{S}$  bereme jako konzistentní odhad matice  $\tilde{\mathbf{\Omega}}$ , podrobněji viz Cipra (1984, odstavec 9.1).

Pro praktickou aplikaci je zpravidla potřebné znát odhady parametrů v původním strukturálním tvaru soustavy, až ty nám totiž odhalí skutečné vztahy mezi jednotlivými proměnnými.

Pokud by se nám podařilo parametry  $\mathbf{\Pi}$  redukováného tvaru transformovat (jednoznačně) zpět na parametry strukturálního tvaru  $\mathbf{\Gamma}$  a  $\mathbf{B}$ , potom shodná transformace vytvoří z konzistentního OLS-odhadu  $\mathbf{P}$  parametru  $\mathbf{\Pi}$  konzistentní odhad  $\hat{\mathbf{\Gamma}}$  a  $\hat{\mathbf{B}}$  parametrů  $\mathbf{\Gamma}$  a  $\mathbf{B}$ . Takový odhad zpravidla označujeme jako *nepřímý odhad metodou nejmenších čtverců* (*ILS-odhad, Indirect Least Squares*).

V této souvislosti nicméně narážíme na problém, se kterým jsme se setkali již v odstavci 1.2 věnovaném identifikaci. Obecně totiž není zaručena existence (jednoznačné) transformace inverzní k  $\mathbf{\Pi} = -\mathbf{B}\mathbf{\Gamma}^{-1}$ . Podrobný rozbor této problematiky uvádí např. Dhrymes (1994, odstavec 4.5), z tohoto pro ILS-odhad parametrů  $j$ -té rovnice soustavy vyplývají následující závěry:

1. Pro neidentifikovanou rovnici neexistuje ILS-odhad.
2. Pro přesně identifikovanou rovnici existuje právě jeden ILS-odhad, ten je konzistentní a rovněž asymptoticky eficientní, neboť se v tomto případě shoduje s 2SLS-odhadem (viz odstavec 1.4.4).
3. Pro přeidentifikovanou rovnici existují aspoň dva odlišné ILS-odhady, přičemž obecně existuje nejvýše

$$\begin{pmatrix} k - k_j \\ m_j - 1 \end{pmatrix}$$

různých ILS-odhadů, každý z nich je konzistentní, ale již není asymptoticky eficientní, protože žádný z možných odhadů plně nevyužije veškerou dostupnou informaci.

Jak již bylo dříve řečeno, většina soustav simultánních rovnic bývá v praxi přeidentifikovaná, takže ILS-odhady nepatří k běžně používaným odhadovým procedurám.

### 1.4.2 Odhad metodou instrumentálních proměnných

Matici  $(T \times (m_j + k_j - 1))$  proměnných  $\mathbf{W}_j$  nazveme maticí *instrumentálních proměnných* (*Instrumental Variables, IV*) příslušejících  $j$ -té rovnici soustavy simultánních rovnic (1.1.2)

$$\mathbf{y}_j = \mathbf{Y}_j \boldsymbol{\gamma}_j + \mathbf{X}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j = \mathbf{Z}_j \boldsymbol{\delta}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, m, \quad (1.4.5)$$

jestliže vyhovuje následujícím požadavkům:

(IV1)  $\text{plim}(1/T)\mathbf{W}'_j\mathbf{Z}_j = \boldsymbol{\Sigma}_{\mathbf{w}_j\mathbf{z}_j}$ ,  $\boldsymbol{\Sigma}_{\mathbf{w}_j\mathbf{z}_j}$  je konečná regulární matice;

(IV2)  $\text{plim}(1/T)\mathbf{W}'_j\boldsymbol{\varepsilon}_j = \mathbf{0}$ .

Vidíme, že (IV1), resp. (IV2) je asymptotickou verzí předpokladu, že IV-proměnné jsou *korelované s původními vysvětlujícími proměnnými*, resp. že jsou *neko-relované s residuální složkou*.

IV-odhad  $\mathbf{d}_j^{IV}$  parametru  $\boldsymbol{\delta}_j$  je následně definován jako

$$\mathbf{d}_j^{IV} = (\mathbf{W}'_j\mathbf{Z}_j)^{-1}\mathbf{W}'_j\mathbf{y}_j. \quad (1.4.6)$$

Konzistentním odhadem asymptotické rozptylové matice  $\mathbf{d}_j^{IV}$  odhadu parametru  $\boldsymbol{\delta}_j$  je

$$\hat{\sigma}_{jj}^{IV} (\mathbf{Z}'_j\mathbf{W}_j(\mathbf{W}'_j\mathbf{W}_j)^{-1}\mathbf{W}'_j\mathbf{Z}_j)^{-1}, \quad \hat{\sigma}_{jj}^{IV} = \frac{(\mathbf{y}_j - \mathbf{Z}_j\mathbf{d}_j^{IV})'(\mathbf{y}_j - \mathbf{Z}_j\mathbf{d}_j^{IV})}{T}, \quad (1.4.7)$$

kde  $\hat{\sigma}_{jj}^{IV}$  je konzistentním odhadem  $\sigma_{jj}$ .

IV-odhad parametrů  $j$ -té rovnice má následující vlastnosti:

- je konzistentní, neboť

$$\text{plim } \mathbf{d}_j^{IV} = \boldsymbol{\delta}_j + \text{plim} \left( \frac{\mathbf{W}'_j\mathbf{Z}_j}{T} \right)^{-1} \text{plim} \left( \frac{\mathbf{W}'_j\boldsymbol{\varepsilon}_j}{T} \right) = \boldsymbol{\delta}_j;$$

- je asymptoticky normální: pro  $T \rightarrow \infty$  platí

$$\sqrt{T} (\mathbf{d}_j^{IV} - \boldsymbol{\delta}_j) \xrightarrow{d} \mathbf{N}(\mathbf{0}, \text{plim } \mathbf{S}_{jT}), \quad (1.4.8)$$

kde

$$\mathbf{S}_{jT} = \hat{\sigma}_{jj}^{IV} \left( \frac{\mathbf{W}'_j \mathbf{Z}_j}{T} \right)^{-1} \left( \frac{\mathbf{W}'_j \mathbf{W}_j}{T} \right) \left( \frac{\mathbf{Z}'_j \mathbf{W}_j}{T} \right)^{-1}, \quad (1.4.9)$$

jestliže posloupnost náhodných veličin  $\varepsilon_{j1}, \varepsilon_{j2}, \dots$  konverguje v distribuci k náhodné veličině s normálním rozdělením; konvergenci v distribuci označujeme jako v (1.4.8) symbolem  $d$ , její definici uvádí Anděl (2007, dodatek B).

Zásadní otázkou konstrukce IV-odhadu je vlastní volba IV-proměnných, zpravidla automaticky se mezi IV-proměnné zařazují exogenní proměnné, neboť jsou nekorelované s residuální složkou.

Kvůli problémům s výběrem vhodných IV-proměnných se v praxi tento odhad v uvedené podobě příliš nepoužívá, jedná se tak spíše o teoretické východisko.

### 1.4.3 Momentové odhady

*Momentové odhady*, speciálně *zobecněný momentový odhad* (*GMM-odhad*, *Generalized Method of Moments*), zastřešují velkou část běžně používaných odhadových procedur. Obvykle mívají výhodné vlastnosti (konzistenci, asymptotickou normalitu a eficienci) bez nutnosti silných předpokladů. V tomto odstavci se seznámíme se základními principy této metody, další podrobnosti lze nalézt v Greene (2003, kapitola 18) nebo Hall (2005).

Momentové odhady jsou založeny na teoretickém vyjádření momentů určitých veličin vyplývajících z odhadovaného modelu a závisejících na parametrech tohoto modelu, zpravidla je zapisujeme ve tvaru

$$\mathbf{E}(\mathbf{m}_t(\boldsymbol{\theta})) = \mathbf{0}, \quad t = 1, \dots, T, \quad (1.4.10)$$

kde  $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^p$  je odhadovaný  $p$ -rozměrný parametr a  $\mathbf{m}_t : \Theta \subseteq \mathbb{R}^p \rightarrow \mathbb{R}^q$ ,  $\infty > q \geq p$ , je vhodné měřitelné zobrazení. Zpravidla se kromě jiného předpokládá, že  $\Theta \subseteq \mathbb{R}^p$  je konvexní či kompaktní množina, skutečná hodnota parametru  $\boldsymbol{\theta}$  leží uvnitř této množiny, dále že  $\mathbf{m}_t$  je spojitě zobrazení, jehož derivace podle  $\boldsymbol{\theta}$  existuje a jedná se o spojitě zobrazení na  $\Theta \subseteq \mathbb{R}^p$ .

Zobecněný momentový odhad je definován vztahem

$$\mathbf{E}(\bar{\mathbf{m}}_T(\boldsymbol{\theta})) = \mathbf{0}, \quad \text{kde } \bar{\mathbf{m}}_T(\boldsymbol{\theta}) = \frac{1}{T} \sum_{t=1}^T \mathbf{m}_t(\boldsymbol{\theta}). \quad (1.4.11)$$

Pro získání odhadu parametru  $\boldsymbol{\theta}$  z podmínek (1.4.10) a (1.4.11) musíme v obou případech využít jejich výběrových verzí.

GMM-odhad parametru  $\boldsymbol{\theta}$  podle (1.4.11) obvykle kvůli možné nejednoznačnosti řešení hledáme pomocí optimalizační úlohy

$$\min_{\boldsymbol{\theta} \in \Theta} \bar{\mathbf{m}}_T(\boldsymbol{\theta})' \mathbf{W}_T \bar{\mathbf{m}}_T(\boldsymbol{\theta}), \quad (1.4.12)$$

kde  $\mathbf{W}_T$  je libovolná pozitivně definitní matice ( $q \times q$ ) nezávislejší na parametru  $\boldsymbol{\theta}$ , u níž se většinou předpokládá, že pro  $T \rightarrow \infty$  konverguje v pravděpodobnosti k pozitivně definitní matici  $\mathbf{W}$  typu ( $q \times q$ ). V případě  $p = q$  se úloha (1.4.12) jednoduše řeší jako soustava  $q$  rovnic o  $p(=q)$  neznámých. Poznamenejme, že si odhad získaný z (1.4.12) zachovává (za platnosti obecných předpokladů) výhodné vlastnosti (konzistenci, asymptotickou normalitu a eficienci).

Podívejme se nyní, jakým způsobem lze GMM-odhad využít v souvislosti se soustavami simultánních rovnic.

Uvažujme soustavu simultánních rovnic ve tvaru (1.1.2)

$$\mathbf{y}_j = \mathbf{Z}_j \boldsymbol{\delta}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, m, \quad (1.4.13)$$

v jednotlivých okamžicích  $t = 1, \dots, T$  tedy platí

$$y_{jt} = \mathbf{z}_{jt}' \boldsymbol{\delta}_j + \varepsilon_{jt}, \quad j = 1, \dots, m, \quad (1.4.14)$$

kde  $\mathbf{z}_{jt}$  označuje  $t$ -tý řádek matice  $\mathbf{Z}_j$  typu ( $T \times (k_j + m_j - 1)$ ).

Z uvedeného předpokladu ortogonality (**P2**) vyplývá požadavek

$$\mathbf{E}(\mathbf{x}'_t \varepsilon_{jt}) = \mathbf{E}(\mathbf{x}'_t (y_{jt} - \mathbf{z}_{jt}' \boldsymbol{\delta}_j)) = \mathbf{0}, \quad j = 1, \dots, m, \quad (1.4.15)$$

kde  $\mathbf{x}_t$  označuje  $t$ -tý řádek matice  $\mathbf{X}$  typu ( $T \times k$ ). Vidíme, že uvedená podmínka je v souladu s (1.4.10), kde  $\mathbf{m}_t(\boldsymbol{\delta}_j) = \mathbf{x}'_t (y_{jt} - \mathbf{z}_{jt}' \boldsymbol{\delta}_j)$ . Zobecněný momentový odhad je v tomto případě zřejmě definován jako

$$\mathbf{E}(\bar{\mathbf{m}}_T(\boldsymbol{\theta})) = \mathbf{0}, \quad \text{kde } \bar{\mathbf{m}}_T(\boldsymbol{\delta}_j) = \frac{1}{T} \sum_{t=1}^T \mathbf{m}_t(\boldsymbol{\delta}_j) = \frac{1}{T} \sum_{t=1}^T \mathbf{x}'_t (y_{jt} - \mathbf{z}_{jt}' \boldsymbol{\delta}_j), \quad (1.4.16)$$

optimalizační úlohu (1.4.12) pro nalezení GMM-odhadu parametru  $\boldsymbol{\delta}_j$  nyní formulujeme ve tvaru

$$\min_{\boldsymbol{\delta}_j} \bar{\mathbf{m}}_T(\boldsymbol{\delta}_j)' \mathbf{W}_{jT} \bar{\mathbf{m}}_T(\boldsymbol{\delta}_j). \quad (1.4.17)$$

Položíme-li derivaci účelové funkce v (1.4.17) podle  $\boldsymbol{\delta}_j$  rovnu nulovému vektoru, s pomocí jednoduchých úprav obdržíme  $\mathbf{d}_j$  odhad parametru  $\boldsymbol{\delta}_j$ ,

$$\mathbf{d}_j = (\mathbf{Z}'_j \mathbf{X} \mathbf{W}_{jT} \mathbf{X}' \mathbf{Z}_j)^{-1} \mathbf{Z}'_j \mathbf{X} \mathbf{W}_{jT} \mathbf{X}' \mathbf{y}_j, \quad (1.4.18)$$

existuje-li inverzní matice k  $\mathbf{Z}'_j \mathbf{X} \mathbf{W}_{jT} \mathbf{X}' \mathbf{Z}_j$ .

**Tvrzení 1.4.1.** *Nechť je  $h(\mathbf{E}(\mathbf{x}'_t \mathbf{z}_{jt})) = h(\mathbf{C}) = k_j + m_j - 1$  a nechť pozitivně definitní matice  $\mathbf{W}_{jT}$  typu ( $k \times k$ ) pro  $T \rightarrow \infty$  konverguje v pravděpodobnosti k pozitivně definitní matici  $\mathbf{W}_j$  typu ( $k \times k$ ), potom je  $\mathbf{d}_j$  konzistentním odhadem parametru  $\boldsymbol{\delta}_j$ .*

*Důkaz.* Jednoduchými úpravami obdržíme

$$\mathbf{d}_j = \boldsymbol{\delta}_j + (\mathbf{Z}'_j \mathbf{X} \mathbf{W}_{jT} \mathbf{X}' \mathbf{Z}_j)^{-1} \mathbf{Z}'_j \mathbf{X} \mathbf{W}_{jT} \mathbf{X}' \boldsymbol{\varepsilon}_j.$$

Člen  $(\mathbf{Z}'_j \mathbf{X} \mathbf{W}_{jT} \mathbf{X}' \mathbf{Z}_j)^{-1} \mathbf{Z}'_j \mathbf{X} \mathbf{W}_{jT} \mathbf{X}' \boldsymbol{\varepsilon}_j$  lze rovněž vyjádřit jako

$$\left[ \left( \frac{1}{T} \sum_{t=1}^T \mathbf{z}'_{jt} \mathbf{x}_t \right) \mathbf{W}_{jT} \left( \frac{1}{T} \sum_{t=1}^T \mathbf{x}'_t \mathbf{z}_{jt} \right) \right]^{-1} \left( \frac{1}{T} \sum_{t=1}^T \mathbf{z}'_{jt} \mathbf{x}_t \right) \mathbf{W}_{jT} \left( \frac{1}{T} \sum_{t=1}^T \mathbf{x}'_t \varepsilon_{jt} \right).$$

Matice  $\mathbf{C} = \mathbf{E}(\mathbf{x}'_t \mathbf{z}_{jt})$  typu  $(k \times (k_j + m_j - 1))$  má podle předpokladů věty plnou sloupcovou hodnotu, odtud zřejmě plyne  $k_j + m_j - 1 \leq k$ , nebo-li  $(k - k_j) + (m - m_j) \geq m - 1$ , což je ve shodě s rozměrovou podmínkou identifikace (1.2.7). Matice  $\mathbf{W}_j$  je navíc pozitivně definitní, a tak je součin  $\mathbf{C}'\mathbf{W}_j\mathbf{C}$  regulární matice.

Podle slabého zákona velkých čísel a předpokladu **(A2)** dostaneme

$$\text{plim } \mathbf{d}_j = \boldsymbol{\delta}_j + (\mathbf{C}'\mathbf{W}_j\mathbf{C})^{-1}\mathbf{C}'\mathbf{W}_j \text{ plim } \frac{1}{T} \sum_{t=1}^T \mathbf{x}'_t \varepsilon_{jt} = \boldsymbol{\delta}_j + (\mathbf{C}'\mathbf{W}_j\mathbf{C})^{-1}\mathbf{C}'\mathbf{W}_j \mathbf{0} = \boldsymbol{\delta}_j.$$

□

**Tvrzení 1.4.2.** *Nechť jsou splněny předpoklady předchozího tvrzení a dále necht' pro  $T \rightarrow \infty$  platí*

$$\bar{\mathbf{m}}_T(\boldsymbol{\delta}_j) = \frac{1}{T} \sum_{t=1}^T \mathbf{x}'_t (y_{jt} - \mathbf{z}_{jt} \boldsymbol{\delta}_j) \xrightarrow{d} \mathbf{N}(\mathbf{0}, \boldsymbol{\Phi}_j/T), \quad (1.4.19)$$

potom pro  $T \rightarrow \infty$  získáváme

$$\sqrt{T}(\mathbf{d}_j - \boldsymbol{\delta}_j) \xrightarrow{d} \mathbf{N}(\mathbf{0}, (\mathbf{C}'\mathbf{W}_j\mathbf{C})^{-1}\mathbf{C}'\mathbf{W}_j\boldsymbol{\Phi}_j\mathbf{W}_j\mathbf{C}(\mathbf{C}'\mathbf{W}_j\mathbf{C})^{-1}). \quad (1.4.20)$$

*Důkaz.* Viz Greene (2003, odstavec 18.3.3). □

**Tvrzení 1.4.3.** *Nechť jsou splněny předpoklady předchozího tvrzení a platí  $\mathbf{W}_j = \boldsymbol{\Phi}_j^{-1}$ , potom je  $\mathbf{d}_j$  asymptoticky eficientním odhadem parametru  $\boldsymbol{\delta}_j$ .*

*Důkaz.* Viz diskuse ve Wooldridge (2002, odstavec 8.3.3). □

V případě, že se nám podaří matici  $\boldsymbol{\Phi}_j$  konzistentně odhadnout maticí  $\boldsymbol{\Phi}_{jT}$  a položíme-li  $\mathbf{W}_{jT} = \boldsymbol{\Phi}_{jT}^{-1}$ , získáme asymptoticky eficientní GMM-odhad parametru  $\boldsymbol{\delta}_j$  s konzistentně odhadnutou asymptotickou rozptylovou maticí

$$(\mathbf{Z}'_j \mathbf{X}(\mathbf{W}_{jT}/T) \mathbf{X}' \mathbf{Z}_j)^{-1}. \quad (1.4.21)$$

**Poznámka 1.4.1.**

1. Optimální volbou matice  $\boldsymbol{\Phi}_j$  ve smyslu asymptotické eficeince GMM-odhadu je podle předchozích tvrzení asymptotická varianční matice  $\sqrt{T}\bar{\mathbf{m}}_T(\boldsymbol{\delta}_j)$ , tj.

$$\boldsymbol{\Phi}_j = \text{asy.var}(\sqrt{T}\bar{\mathbf{m}}_T(\boldsymbol{\delta}_j)) = \text{plim } \frac{1}{T} \sum_{t=1}^T \mathbf{x}'_t \mathbf{x}_t (y_{jt} - \mathbf{z}_{jt} \boldsymbol{\delta}_j)^2, \quad (1.4.22)$$

kteřá se v praxi konzistentně odhaduje pomocí

$$\boldsymbol{\Phi}_{jT} = \frac{1}{T} \sum_{t=1}^T \mathbf{x}'_t \mathbf{x}_t (y_{jt} - \mathbf{z}_{jt} \mathbf{d}_j)^2, \quad (1.4.23)$$

kde  $\mathbf{d}_j$  označuje konzistentní odhad  $\boldsymbol{\delta}_j$ .

2. Jestliže jsou residuální složky  $\varepsilon_{jt}$  v  $j$ -té rovnici homoskedastické, tj.  $\text{var}(\varepsilon_{jt}) = \sigma_{jj}$  pro všechna  $t$  a  $\text{cov}(\varepsilon_{js}, \varepsilon_{jt}) = 0$  pro všechna  $s \neq t$ , což v soustavách simultánních rovnic apriori předpokládáme, potom nejvhodnější volbou matice  $\Phi_j$  je

$$\Phi_j = \text{plim} \frac{\sigma_{jj} \mathbf{X}'\mathbf{X}}{T}. \quad (1.4.24)$$

Konstanta  $\sigma_{jj}$  není pro řešení úlohy (1.4.17) vzhledem k (1.4.18) podstatná, obvykle tak užíváme  $\mathbf{W}_{jT} = (\Phi_{jT})^{-1} = (\hat{\sigma}_{jj} \mathbf{X}'\mathbf{X}/T)^{-1}$ , přičemž  $\hat{\sigma}_{jj}$  je konzistentním odhadem parametru  $\sigma_{jj}$ . Konečná podoba odhadu  $\mathbf{d}_j$  parametru  $\delta_j$  je

$$\mathbf{d}_j^{2SLS} = (\mathbf{Z}'_j \mathbf{X} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{Z}_j)^{-1} \mathbf{Z}'_j \mathbf{X} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{y}_j. \quad (1.4.25)$$

Předchozí odhad se nazývá *dvoustupňový odhad metodou nejmenších čtverců* (*2SLS-odhad*, *Two Stage Least Squares*), podrobněji se s ním seznámíme v odstavci 1.4.4.

3. Pokud bychom požadavek homoskedasticity residuí v  $j$ -té rovnici z nějakého důvodu ignorovali, Greene (2003, odstavec 15.5.4) doporučuje použít tzv. *H2SLS-odhad* (*Heteroscedastic 2SLS*)

$$\mathbf{d}_j^{H2SLS} = (\mathbf{Z}'_j \mathbf{X} \mathbf{S}_{jT}^{-1} \mathbf{X}' \mathbf{Z}_j)^{-1} \mathbf{Z}'_j \mathbf{X} \mathbf{S}_{jT}^{-1} \mathbf{X}' \mathbf{y}_j, \quad (1.4.26)$$

kde

$$\mathbf{S}_{jT} = \frac{1}{T} \sum_{t=1}^T \mathbf{x}'_t \mathbf{x}_t (y_{jt} - \mathbf{z}_{jt} \mathbf{d}_j^{2SLS})^2 \quad [= \Phi_{jT}], \quad (1.4.27)$$

neboť 2SLS-odhad by byl v tomto případě konzistentní, ale nebyl by eficientní, právě tento nedostatek odstraňuje zmíněný H2SLS-odhad.

◇

**Poznámka 1.4.2.** V kontextu studované problematiky se v literatuře, např. ve Wooldridge (2002, odstavec 8.2.5), nezřídka setkáváme s testem hypotézy tzv. *over-identification restrictions*, při kterém ověřujeme platnost nulové hypotézy, která tvrdí, že exogenní proměnné a residuální složka  $j$ -té rovnice soustavy jsou současně nekorelované, tj. je splněna podmínka ortogonality, proti alternativě, že tomu tak není. Příslušná testová statistika

$$J_j = T \bar{\mathbf{m}}_T(\hat{\delta}_j)' \Phi_{jT}^{-1} \bar{\mathbf{m}}_T(\hat{\delta}_j) \quad (1.4.28)$$

má za platnosti nulové hypotézy asymptoticky rozdělení  $\chi^2((k - k_j) - (m_j - 1))$ . Poznamenejme, že s uvedenou statistikou lze pracovat pouze pro  $k > k_j + m_j - 1$ .

Zamítnutí nulové hypotézy implikuje, že některá z exogenních proměnných nesplňuje požadavek ortogonality s residuální složkou. To může být způsobeno buď tím, že se mezi exogenními proměnnými nachází taková, která k nim fakticky nenáleží, nebo tím, že nějaká exogenní proměnná byla z  $j$ -té rovnice nevhodně vynechána.



Položíme-li speciálně

$$\Phi_{jT} = \hat{\sigma}_{jj} (\mathbf{X}'\mathbf{X}/T), \quad \hat{\sigma}_{jj} = \frac{\hat{\boldsymbol{\varepsilon}}_j' \hat{\boldsymbol{\varepsilon}}_j}{T}, \quad \hat{\boldsymbol{\varepsilon}}_j = \mathbf{y}_j - \mathbf{Z}_j \hat{\boldsymbol{\delta}}_j \quad \text{a} \quad \hat{\boldsymbol{\delta}}_j = \mathbf{d}_j^{2SLS},$$

úpravou (1.4.28) obdržíme *Sarganovu statistiku*

$$Sa_j = \frac{\hat{\boldsymbol{\varepsilon}}_j' \mathbf{X} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \hat{\boldsymbol{\varepsilon}}_j}{\hat{\boldsymbol{\varepsilon}}_j' \hat{\boldsymbol{\varepsilon}}_j / T}. \quad (1.4.29)$$

Není obtížné nahlédnout, že  $Sa_j = T \cdot \tilde{R}_j^2$ , kde  $\tilde{R}_j^2$  je tzv. *necentrováný koeficient determinace* v regresi  $\hat{\boldsymbol{\varepsilon}}_j$  na všechny exogenní proměnné modelu. Uvedme, že necentrováný koeficient je v lineárním regresním modelu  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ ,  $\boldsymbol{\varepsilon} \sim (\mathbf{0}, \boldsymbol{\sigma}^2 \mathbf{I})$ , definován jako

$$\tilde{R}^2 = 1 - \frac{(\mathbf{y} - \hat{\mathbf{y}})'(\mathbf{y} - \hat{\mathbf{y}})}{\hat{\mathbf{y}}'\hat{\mathbf{y}}}.$$

Tudíž (asymptotický) kritický obor pro test prezentované hypotézy na hladině významnosti  $\alpha$  je v tomto specifickém případě tvaru

$$Sa_j = T \cdot \tilde{R}_j^2 \geq \chi_{1-\alpha}^2 ((k - k_j) - (m_j - 1)). \quad (1.4.30)$$

◇

V předcházejícím textu jsme odvodili některé vybrané postupy odhadu parametrů v jednotlivých rovnicích soustavy simultánních rovnic. Nyní se budeme snažit efektivněji využít informace, které jsou ukryty ve vztazích mezi jednotlivými rovnicemi, tj. vzájemnou korelovanost jejich residuálních složek.

Nejprve všech  $m$  rovnic strukturálního tvaru soustavy (1.1.2)

$$\mathbf{y}_j = \mathbf{Y}_j \boldsymbol{\gamma}_j + \mathbf{X}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j = \mathbf{Z}_j \boldsymbol{\delta}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, m, \quad (1.4.31)$$

vyjádříme jediným modelem

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{Z}_m \end{pmatrix} \begin{pmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \\ \vdots \\ \boldsymbol{\delta}_m \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_m \end{pmatrix} = \mathbf{Z}^* \boldsymbol{\delta} + \boldsymbol{\varepsilon}. \quad (1.4.32)$$

Znovu vycházejme z podmínky (1.4.15), nyní ovšem zformulujeme optimalizační úlohu pro GMM-odhad na základě znalosti vztahů v celé soustavě rovnic. Položme  $\bar{\mathbf{m}}_T(\boldsymbol{\delta}) = (\bar{\mathbf{m}}_T(\boldsymbol{\delta}_1)', \dots, \bar{\mathbf{m}}_T(\boldsymbol{\delta}_m)')$ , následně hledáme

$$\min_{\boldsymbol{\delta}} \bar{\mathbf{m}}_T(\boldsymbol{\delta})' \mathbf{W}_T \bar{\mathbf{m}}_T(\boldsymbol{\delta}), \quad (1.4.33)$$

neboli

$$\min_{\boldsymbol{\delta}_j, j=1, \dots, m} \sum_{j=1}^m \sum_{l=1}^m \bar{\mathbf{m}}_T(\boldsymbol{\delta}_j)' \mathbf{W}_{j,l,T} \bar{\mathbf{m}}_T(\boldsymbol{\delta}_l), \quad (1.4.34)$$

kde matice  $\mathbf{W}_{jl,T}$  typu  $(k \times k)$  vystupuje jako  $jl$ -tý blok matice  $\mathbf{W}_T$  typu  $(mk \times mk)$ ,  $j, l = 1, \dots, m$ .

Účelovou funkci ve vztahu (1.4.34) můžeme explicitně vyjádřit jako

$$\begin{pmatrix} \mathbf{X}'(\mathbf{y}_1 - \mathbf{Z}_1\boldsymbol{\delta}_1)/T \\ \mathbf{X}'(\mathbf{y}_2 - \mathbf{Z}_2\boldsymbol{\delta}_2)/T \\ \vdots \\ \mathbf{X}'(\mathbf{y}_m - \mathbf{Z}_m\boldsymbol{\delta}_m)/T \end{pmatrix}' \begin{pmatrix} \mathbf{W}_{11,T} & \dots & \mathbf{W}_{1m,T} \\ \mathbf{W}_{21,T} & \dots & \mathbf{W}_{2m,T} \\ \vdots & \dots & \vdots \\ \mathbf{W}_{m1,T} & \dots & \mathbf{W}_{mm,T} \end{pmatrix} \begin{pmatrix} \mathbf{X}'(\mathbf{y}_1 - \mathbf{Z}_1\boldsymbol{\delta}_1)/T \\ \mathbf{X}'(\mathbf{y}_2 - \mathbf{Z}_2\boldsymbol{\delta}_2)/T \\ \vdots \\ \mathbf{X}'(\mathbf{y}_m - \mathbf{Z}_m\boldsymbol{\delta}_m)/T \end{pmatrix}. \quad (1.4.35)$$

Postupně pro  $j = 1, \dots, m$  položme

$$\frac{\partial}{\partial \boldsymbol{\delta}_j} \sum_{j=1}^m \sum_{l=1}^m \bar{\mathbf{m}}_T(\boldsymbol{\delta}_j)' \mathbf{W}_{jl,T} \bar{\mathbf{m}}_T(\boldsymbol{\delta}_l) = 2 \sum_{l=1}^m \left( \frac{\mathbf{Z}'_j \mathbf{X}}{T} \right) \mathbf{W}_{jl,T} \left[ \frac{\mathbf{X}'(\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\delta}_l)}{T} \right] = \mathbf{0}. \quad (1.4.36)$$

Při obdobných předpokladech jako výše si takový GMM-odhad zachovává zmiňované vlastnosti (konzistenci, asymptotickou normalitu a eficienci), tentokrát ovšem v rámci celé soustavy a nikoliv pouze pro jednotlivé rovnice.

Speciálně optimálně volenou maticí  $\mathbf{W}$  (pozitivně definitní matice  $\mathbf{W}_T$  konverguje v pravděpodobnosti pro  $T \rightarrow \infty$  k pozitivně definitní matici  $\mathbf{W}$ ) ve smyslu asymptotické eficeie bude opět  $\mathbf{W} = \boldsymbol{\Phi}^{-1}$ , přičemž  $\boldsymbol{\Phi}$  je asymptotická varianční matice  $\sqrt{T}\bar{\mathbf{m}}_T(\boldsymbol{\delta})$  skládající se z matic  $\boldsymbol{\Phi}_{jl}$ . Detailněji nahlédneme

$$\boldsymbol{\Phi}_{jl} = \text{plim} \left( \frac{1}{T} \sum_{t=1}^T \mathbf{x}'_t \mathbf{x}_t (y_{jt} - z_{jt} \boldsymbol{\delta}_j)(y_{lt} - z_{lt} \boldsymbol{\delta}_l) \right). \quad (1.4.37)$$

V praktických výpočtech se matice  $\boldsymbol{\Phi}_{jl}$  konzistentně odhadují pomocí

$$\hat{\boldsymbol{\Phi}}_{jl} = \frac{1}{T} \sum_{t=1}^T \mathbf{x}'_t \mathbf{x}_t (y_{jt} - z_{jt} \mathbf{d}_j)(y_{lt} - z_{lt} \mathbf{d}_l), \quad (1.4.38)$$

kde  $\mathbf{d}_j$ ,  $j = 1, \dots, m$ , je konzistentní odhad parametru  $\boldsymbol{\delta}_j$  (zpravidla 2SLS-odhad).

Greene (2003, odstavec 15.6.3) jako řešení všech  $m$  rovnic (1.4.36) s přihlédnutím k (1.4.38) uvádí

$$\begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_m \end{pmatrix} = \mathbf{KL}, \quad (1.4.39)$$

kde

$$\mathbf{K} = \begin{pmatrix} \mathbf{Z}'_1 \mathbf{X} \hat{\boldsymbol{\Phi}}^{11} \mathbf{X}' \mathbf{Z}_1 & \mathbf{Z}'_1 \mathbf{X} \hat{\boldsymbol{\Phi}}^{12} \mathbf{X}' \mathbf{Z}_2 & \dots & \mathbf{Z}'_1 \mathbf{X} \hat{\boldsymbol{\Phi}}^{1m} \mathbf{X}' \mathbf{Z}_m \\ \mathbf{Z}'_2 \mathbf{X} \hat{\boldsymbol{\Phi}}^{21} \mathbf{X}' \mathbf{Z}_1 & \mathbf{Z}'_2 \mathbf{X} \hat{\boldsymbol{\Phi}}^{22} \mathbf{X}' \mathbf{Z}_2 & \dots & \mathbf{Z}'_2 \mathbf{X} \hat{\boldsymbol{\Phi}}^{2m} \mathbf{X}' \mathbf{Z}_m \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{Z}'_m \mathbf{X} \hat{\boldsymbol{\Phi}}^{m1} \mathbf{X}' \mathbf{Z}_1 & \mathbf{Z}'_m \mathbf{X} \hat{\boldsymbol{\Phi}}^{m2} \mathbf{X}' \mathbf{Z}_2 & \dots & \mathbf{Z}'_m \mathbf{X} \hat{\boldsymbol{\Phi}}^{mm} \mathbf{X}' \mathbf{Z}_m \end{pmatrix}^{-1}, \quad (1.4.40)$$

$$L = \begin{pmatrix} \sum_{j=1}^m \mathbf{Z}'_1 \mathbf{X} \hat{\Phi}^{1j} \mathbf{y}_j \\ \sum_{j=1}^m \mathbf{Z}'_2 \mathbf{X} \hat{\Phi}^{2j} \mathbf{y}_j \\ \vdots \\ \sum_{j=1}^m \mathbf{Z}'_m \mathbf{X} \hat{\Phi}^{mj} \mathbf{y}_j \end{pmatrix}, \quad (1.4.41)$$

kde  $\hat{\Phi}^{jl}$  je  $jl$ -tý blok matice  $\hat{\Phi}^{-1}$ .

### Poznámka 1.4.3.

1. Jestliže  $\hat{\Phi}_{jj} = \hat{\sigma}_{jj}(\mathbf{X}'\mathbf{X}/T)$  a  $\hat{\Phi}_{jl} = 0$  pro  $j \neq l$ , potom  $\mathbf{d}_j$  je 2SLS-odhadem.
2. Jestliže  $\hat{\Phi}_{jl} = 0$  pro  $j \neq l$ , potom  $\mathbf{d}_j$  je H2SLS-odhadem.
3. Jestliže  $\hat{\Phi}_{jl} = \hat{\sigma}_{jl}(\mathbf{X}'\mathbf{X}/T)$ , potom  $\mathbf{d}_j$  je tzv. *třístupňovým odhadem metodou nejmenších čtverců (3SLS-odhad, Three Stage Least Squares)*, kterému se budeme podrobněji věnovat v odstavci 1.4.5.

◇

Na závěr uveďme, že pro asymptoticky eficientní GMM-odhad  $\mathbf{d}$  parametru  $\boldsymbol{\delta}$  za analogických předpokladů, které byly uvedeny v rámci tvrzení prezentovaných v tomto odstavci, pro  $T \rightarrow \infty$  získáváme

$$\sqrt{T}(\mathbf{d} - \boldsymbol{\delta}) \xrightarrow{d} \mathbf{N}(\mathbf{0}, \text{plim } \mathbf{S}_T), \quad (1.4.42)$$

kde

$$\mathbf{S}_T = \left[ \frac{\mathbf{Z}'^* (\mathbf{I}_m \otimes \mathbf{X})}{T} (\hat{\Phi}^{-1}) \frac{(\mathbf{I}_m \otimes \mathbf{X})' \mathbf{Z}^*}{T} \right]^{-1}. \quad (1.4.43)$$

### 1.4.4 2SLS-odhad

Dvoustupňový odhad metodou nejmenších čtverců zmiňovaný v předchozím odstavci dnes zřejmě patří k nejpoužívanějším odhadovým procedurám v rámci teorie soustav simultánních rovnic, poskytuje totiž prakticky přijatelné výsledky.

Konstrukce 2SLS-odhadu se obvykle provádí ve dvou stupních. Uvažujme soustavu simultánní rovnic ve tvaru (1.1.2)

$$\mathbf{y}_j = \mathbf{Y}_j \boldsymbol{\gamma}_j + \mathbf{X}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j = \mathbf{Z}_j \boldsymbol{\delta}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, m. \quad (1.4.44)$$

V  $j$ -té rovnici soustavy se odhad příslušných parametrů  $\boldsymbol{\gamma}_j$  a  $\boldsymbol{\beta}_j$  provádí následovně:

- v prvním stupni se získají vypočtené OLS-hodnoty všech endogenních proměnných  $\mathbf{Y}_j$  na pravé straně  $j$ -té rovnice pomocí lineární regrese těchto proměnných na všechny exogenní proměnné  $\mathbf{X}$  soustavy, tj.

$$\hat{\mathbf{Y}}_j = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}_j, \quad (1.4.45)$$

- ve druhém stupni se vypočtou finální verze odhadů  $\mathbf{c}_j^{2SLS}$  a  $\mathbf{b}_j^{2SLS}$  parametrů  $\boldsymbol{\gamma}_j$  a  $\boldsymbol{\beta}_j$  tak, že se v původní rovnici (1.4.44) nahradí regresory  $\mathbf{Y}_j$  vypočtenými hodnotami  $\hat{\mathbf{Y}}_j$ , tj.

$$\mathbf{y}_j = \hat{\mathbf{Y}}_j \boldsymbol{\gamma}_j + \mathbf{X}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j, \quad (1.4.46)$$

a opět se zkonstruuje OLS-odhad v lineárním regresním modelu (1.4.46).

Poznamenejme, že 2SLS-odhad lze explicitně vyjádřit vzorcem

$$\begin{pmatrix} \mathbf{c}_j^{2SLS} \\ \mathbf{b}_j^{2SLS} \end{pmatrix} = \begin{pmatrix} \mathbf{Y}'_j \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y}_j & \mathbf{Y}'_j \mathbf{X}_j \\ \mathbf{X}'_j \mathbf{Y}_j & \mathbf{X}'_j \mathbf{X}_j \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{Y}'_j \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{y}_j \\ \mathbf{X}'_j \mathbf{y}_j \end{pmatrix}, \quad (1.4.47)$$

přičemž jako konzistentní odhad asymptotické rozptylové matice odhadu  $\mathbf{d}_j^{2SLS}$  zpravidla uvažujeme

$$\hat{\sigma}_{jj}^{2SLS} \begin{pmatrix} \mathbf{Y}'_j \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y}_j & \mathbf{Y}'_j \mathbf{X}_j \\ \mathbf{X}'_j \mathbf{Y}_j & \mathbf{X}'_j \mathbf{X}_j \end{pmatrix}^{-1}, \quad (1.4.48)$$

kde

$$\hat{\sigma}_{jj}^{2SLS} = \frac{(\mathbf{y}_j - \mathbf{Y}_j \mathbf{c}_j^{2SLS} - \mathbf{X}_j \mathbf{b}_j^{2SLS})' (\mathbf{y}_j - \mathbf{Y}_j \mathbf{c}_j^{2SLS} - \mathbf{X}_j \mathbf{b}_j^{2SLS})}{T} \quad (1.4.49)$$

je konzistentním odhadem  $\sigma_{jj}$ , podrobnosti nalezneme v Cipra (1984, kapitola 10).

2SLS-odhad parametrů  $j$ -té rovnice má (za platnosti obvyklých asymptotických předpokladů, které jsme uvedli v předchozím odstavci 1.4.3) následující vlastnosti:

- je konzistentní,
- je asymptoticky normální,
- je asymptoticky eficientní v rámci  $j$ -té rovnice soustavy.

Všechny tři vlastnosti byly odvozeny pro GMM-odhad  $j$ -té rovnice soustavy simultánních rovnic, připomeňme, že 2SLS-odhad je speciálním případem tohoto GMM-odhadu.

**Poznámka 1.4.4.** Je také možné ukázat, že 2SLS-odhad je totožný s IV-odhadem z odstavce 1.4.2, jestliže jako matici instrumentálních proměnných vezmeme  $\mathbf{W}_j = (\hat{\mathbf{Y}}_j, \mathbf{X}_j)$ , podrobněji viz Cipra (1984, věta 10.1).  $\diamond$

### 1.4.5 3SLS-odhad

Třístupňový odhad metodou nejmenších čtverců byl stručně představen v odstavci věnovaném GMM-odhadům jako speciální případ odhadového postupu, který lépe využíval dostupnou informaci o ekonometrické soustavě simultánních rovnic.

Již jsme uvedli, že 2SLS-odhad je asymptoticky eficientní v rámci jednotlivých rovnic soustavy, zohledníme-li navíc současnou korelovanost mezi residuálními složkami jednotlivých rovnic, je možné přidáním dalšího stupně obdržet 3SLS-odhad, který bude asymptoticky eficientní v rámci celé soustavy.

Při konstrukci 3SLS-odhadu se zpravidla postupuje v následujících krocích.

V prvním a druhém stupni v souladu s postupem prezentovaným v předchozím odstavci získáme 2SLS-residua pro všechny rovnice soustavy, s jejich pomocí pak konzistentně odhadneme všechny kovariance  $\sigma_{ij}$  jako

$$\hat{\sigma}_{ij} = \frac{(\mathbf{y}_i - \mathbf{Z}_i \mathbf{d}_i^{2SLS})' (\mathbf{y}_j - \mathbf{Z}_j \mathbf{d}_j^{2SLS})}{T}, \quad i, j = 1, \dots, m. \quad (1.4.50)$$

Ve třetím stupni spočteme finální 3SLS-odhad: Opět uvažujme vyjádření všech  $m$  rovnic strukturálního tvaru soustavy (1.1.2) jediným modelem (viz (1.4.32))

$$\mathbf{y} = \mathbf{Z}^* \boldsymbol{\delta} + \boldsymbol{\varepsilon}, \quad \text{var}(\boldsymbol{\varepsilon}) = \tilde{\boldsymbol{\Sigma}} \otimes \mathbf{I}_T = \boldsymbol{\Sigma}, \quad (1.4.51)$$

předpokládáme, že matice  $\tilde{\boldsymbol{\Sigma}}$  je regulární.

Model (1.4.51) vynásobíme zleva maticí  $\mathbf{M}' = (\mathbf{I}_m \otimes \mathbf{X})'$ , tj.

$$\mathbf{M}' \mathbf{y} = \mathbf{M}' \mathbf{Z}^* \boldsymbol{\delta} + \mathbf{M}' \boldsymbol{\varepsilon}. \quad (1.4.52)$$

Tato transformace se provádí kvůli zachování konzistence výsledného odhadu.

Podle Cipra (1984, odstavec 11.1) přibližně platí

$$\text{var}(\mathbf{M}' \boldsymbol{\varepsilon}) \sim \mathbf{M}' \mathbf{E}(\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}') \mathbf{M} = (\mathbf{I}_m \otimes \mathbf{X})' (\tilde{\boldsymbol{\Sigma}} \otimes \mathbf{I}_T) (\mathbf{I}_m \otimes \mathbf{X}) = \tilde{\boldsymbol{\Sigma}} \otimes \mathbf{X}' \mathbf{X}, \quad (1.4.53)$$

rovnost v tomto případě nastává jen za podmínky stochastické nezávislosti  $\mathbf{X}$  a  $\boldsymbol{\varepsilon}$ , zatímco podle předpokladu **(P2)**, resp. **(A2)** máme zaručenu jen jejich současnou (asymptotickou) nekorelovanost. Uvedený vztah je tudíž nutné brát jako hrubě heuristický.

3SLS-odhad  $\mathbf{d}^{3SLS}$  parametru  $\boldsymbol{\delta}$  spočteme jako přípustný Aitkenův odhad (více viz Cipra (2008, odstavec 4.1)) v transformovaném modelu (1.4.52),

$$\mathbf{d}^{3SLS} = \left[ \mathbf{Z}^{*'} \left( \hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \right) \mathbf{Z}^* \right]^{-1} \mathbf{Z}^{*'} \left( \hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \right) \mathbf{y}, \quad (1.4.54)$$

kde

$$\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \dots & \hat{\sigma}_{1m} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} & \dots & \hat{\sigma}_{2m} \\ \vdots & \vdots & \dots & \vdots \\ \hat{\sigma}_{m1} & \hat{\sigma}_{m2} & \dots & \hat{\sigma}_{mm} \end{pmatrix}. \quad (1.4.55)$$

Konzistentním odhadem asymptotické rozptylové matice odhadu  $\mathbf{d}^{3SLS}$  je

$$\left[ \mathbf{Z}^{*'} \left( \hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \right) \mathbf{Z}^* \right]^{-1}. \quad (1.4.56)$$

3SLS-odhad parametrů soustavy simultánních rovnic má (za platnosti obvyklých asymptotických předpokladů) následující vlastnosti:

- je konzistentní,
- je asymptoticky normální (viz poznámka 1.4.3 a závěr odstavce 1.4.3),
- je asymptoticky eficientní v rámci celé soustavy,
- je asymptoticky eficientní vůči 2SLS-odhadu (je-li  $\tilde{\boldsymbol{\Sigma}}$  diagonální nebo jsou-li všechny rovnice soustavy přesně identifikované, jsou oba odhady totožné).

Na závěr poznamenejme, že 3SLS-odhad je speciálním případem GMM-odhadu konstruovaného v rámci celé soustavy simultánních rovnic, viz odstavec 1.4.3.

### 1.4.6 Maximálně věrohodné odhady

V tomto odstavci uvedeme další typy odhadů parametrů v rámci soustav simultánních rovnic, tyto jsou založeny na metodě maximální věrohodnosti; samozřejmě zde apriori předpokládáme znalost příslušného pravděpodobnostního rozdělení residuálních složek, zpravidla uvažujeme (asymptotické) normální rozdělení.

Jako první se seznámíme s *FIML-odhadem* (*Full Information Maximum Likelihood*), se kterým se v praxi i v rámci softwarových produktů setkáme častěji.

Uvažujme strukturální tvar soustavy simultánních rovnic (1.1.3)

$$\mathbf{Y}\boldsymbol{\Gamma} + \mathbf{X}\mathbf{B} + \mathbf{E} = \mathbf{0} \quad , \text{ tj.} \quad (1.4.57)$$

$$\mathbf{y}_t \boldsymbol{\Gamma} + \mathbf{x}_t \mathbf{B} + \boldsymbol{\varepsilon}_t = \mathbf{0}, \quad t = 1, \dots, T, \quad (1.4.58)$$

ve kterém mají náhodné vektory  $\boldsymbol{\varepsilon}'_t$  mnohorozměrné normální rozdělení se střední hodnotou  $\mathbf{E}(\boldsymbol{\varepsilon}'_t) = \mathbf{0}$  a regulární rozptylovou maticí  $\text{var}(\boldsymbol{\varepsilon}'_t) = \tilde{\boldsymbol{\Sigma}}$ .

Věrohodnostní funkce  $T$  nekorelovaných náhodných vektorů  $\boldsymbol{\varepsilon}'_t$  je tvaru

$$L_{\boldsymbol{\varepsilon}} = \prod_{t=1}^T \mathbf{f}(\boldsymbol{\varepsilon}'_t) = \prod_{t=1}^T (2\pi)^{-m/2} |\tilde{\boldsymbol{\Sigma}}|^{-1/2} \exp \left\{ -\frac{\boldsymbol{\varepsilon}_t \tilde{\boldsymbol{\Sigma}}^{-1} \boldsymbol{\varepsilon}'_t}{2} \right\}, \quad (1.4.59)$$

kde  $\mathbf{f}(\cdot)$  značí hustotu mnohorozměrného normálního rozdělení  $\mathbf{N}(\mathbf{0}, \tilde{\boldsymbol{\Sigma}})$ .

Podle (1.4.58) platí

$$\boldsymbol{\varepsilon}_t = -\mathbf{y}_t \boldsymbol{\Gamma} - \mathbf{x}_t \mathbf{B}, \quad t = 1, \dots, T,$$

použijeme-li větu o transformaci hustoty náhodných vektorů (viz Anděl (2007, odstavec 3.2)), získáme věrohodnostní funkci  $T$  náhodných vektorů  $\mathbf{y}'_t$  (jakobiánem příslušné transformace je  $\det \boldsymbol{\Gamma}$ )

$$L = (2\pi)^{-mT/2} (\det \tilde{\boldsymbol{\Sigma}})^{-T/2} |\det \boldsymbol{\Gamma}|^T \exp \left\{ -\frac{1}{2} \sum_{t=1}^T (\mathbf{y}_t \boldsymbol{\Gamma} + \mathbf{x}_t \mathbf{B}) \tilde{\boldsymbol{\Sigma}}^{-1} (\mathbf{y}_t \boldsymbol{\Gamma} + \mathbf{x}_t \mathbf{B})' \right\}. \quad (1.4.60)$$

Poslední člen tohoto vztahu lze s pomocí vlastností stopy matice upravit jako

$$\begin{aligned} \sum_{t=1}^T (\mathbf{y}_t \boldsymbol{\Gamma} + \mathbf{x}_t \mathbf{B}) \tilde{\boldsymbol{\Sigma}}^{-1} (\mathbf{y}_t \boldsymbol{\Gamma} + \mathbf{x}_t \mathbf{B})' &= \sum_{t=1}^T \text{tr} \left( (\mathbf{y}_t \boldsymbol{\Gamma} + \mathbf{x}_t \mathbf{B}) \tilde{\boldsymbol{\Sigma}}^{-1} (\mathbf{y}_t \boldsymbol{\Gamma} + \mathbf{x}_t \mathbf{B})' \right) = \\ &= \text{tr} \left( \tilde{\boldsymbol{\Sigma}}^{-1} \sum_{t=1}^T (\mathbf{y}_t \boldsymbol{\Gamma} + \mathbf{x}_t \mathbf{B})' (\mathbf{y}_t \boldsymbol{\Gamma} + \mathbf{x}_t \mathbf{B}) \right) = T \text{tr} \left( \tilde{\boldsymbol{\Sigma}}^{-1} \mathbf{W} \right), \end{aligned}$$

kde pro matici  $\mathbf{W}$  platí

$$\mathbf{W} = \frac{1}{T} (\mathbf{Y}\boldsymbol{\Gamma} + \mathbf{X}\mathbf{B})' (\mathbf{Y}\boldsymbol{\Gamma} + \mathbf{X}\mathbf{B}).$$

Zlogaritmováním funkce  $L$  v (1.4.60) získáme logaritmickou věrohodnostní funkci

$$l = \log L = -\frac{mT}{2} \log(2\pi) - \frac{T}{2} \log(\det \tilde{\boldsymbol{\Sigma}}) + T \log |\det \boldsymbol{\Gamma}| - \frac{T}{2} \text{tr} \left( \tilde{\boldsymbol{\Sigma}}^{-1} \mathbf{W} \right). \quad (1.4.61)$$

Derivováním (1.4.61) podle prvků matice  $\tilde{\Sigma}^{-1}$  obdržíme metodou maximální věrohodnosti  $\hat{\tilde{\Sigma}}$  odhad  $\tilde{\Sigma}$  v závislosti na neznámých parametrech  $\Gamma$  a  $B$ ,

$$\hat{\tilde{\Sigma}} = \frac{1}{T}(\mathbf{Y}\Gamma + \mathbf{X}B)'(\mathbf{Y}\Gamma + \mathbf{X}B) = \mathbf{W}. \quad (1.4.62)$$

Dosadíme-li (1.4.62) do (1.4.61), získáme tzv. *koncentrovanou logaritmickou věrohodnostní funkci*  $l^*$ ,

$$l^* = -\frac{mT}{2}[\log(2\pi) + 1] + T \log |\det \Gamma| - \frac{T}{2} \log (\det \mathbf{W}). \quad (1.4.63)$$

FIML-odhad parametrů  $\Gamma$  a  $B$  bychom obdrželi maximalizací (1.4.63), nicméně zmíněná funkce je v uvedených parametrech nelineární, řešení proto nelze vyjádřit explicitně a je nutné jej hledat iteračními metodami.

Jednou ze známějších iteračních metod je *linearizovaná maximální věrohodnost*. Srovnáme nejprve prvky matic  $B$  a  $\Gamma$  do jediného vektoru  $\delta$  analogicky jako v (1.4.32).

Požadujeme, aby  $\hat{\delta}$  FIML-odhad parametru  $\delta$  splňoval

$$\frac{\partial l^*}{\partial \delta}(\hat{\delta}) = \mathbf{0}. \quad (1.4.64)$$

Rozvíňme nyní  $\frac{\partial l^*}{\partial \delta}$  v Taylorovu řadu kolem skutečné hodnoty parametru  $\delta^0$ , tj.

$$\frac{\partial l^*}{\partial \delta}(\hat{\delta}) = \frac{\partial l^*}{\partial \delta}(\delta^0) + \frac{\partial^2 l^*}{\partial \delta \partial \delta}(\bar{\delta}) \cdot (\hat{\delta} - \delta^0),$$

kde  $\bar{\delta}$  leží na úsečce mezi  $\hat{\delta}$  a  $\delta^0$ . Na základě podmínky (1.4.64) získáváme

$$\hat{\delta} = \delta^0 - \left[ \frac{\partial^2 l^*}{\partial \delta \partial \delta}(\bar{\delta}) \right]^{-1} \frac{\partial l^*}{\partial \delta}(\delta^0), \quad (1.4.65)$$

předpokládáme-li regularitu matice  $\frac{\partial^2 l^*}{\partial \delta \partial \delta}(\bar{\delta})$ .

Právě prezentovaný vztah (1.4.65) použijeme k iteračnímu výpočtu hodnot  $\delta^{(k)}$ ,  $k \in \mathbb{N}$ ,

$$\delta^{(k)} = \delta^{(k-1)} - \left[ \frac{\partial^2 l^*}{\partial \delta \partial \delta}(\delta^{(k-1)}) \right]^{-1} \frac{\partial l^*}{\partial \delta}(\delta^{(k-1)}), \quad (1.4.66)$$

kde opět předpokládáme regularitu potřebných matic. Výpočet ukončíme, když  $\delta^{(k)} \approx \delta^{(k-1)}$ , pak zřejmě platí

$$\mathbf{0} = \delta^{(k)} - \delta^{(k-1)} = - \left[ \frac{\partial^2 l^*}{\partial \delta \partial \delta}(\delta^{(k-1)}) \right]^{-1} \frac{\partial l^*}{\partial \delta}(\delta^{(k-1)}),$$

tedy pro  $\hat{\delta} = \delta^{(k-1)}$  obdržíme

$$\frac{\partial l^*}{\partial \delta}(\hat{\delta}) = \mathbf{0}.$$

Jako počáteční hodnotu  $\boldsymbol{\delta}^{(0)}$  volíme konzistentní odhad parametru  $\boldsymbol{\delta}^0$  (zpravidla 2SLS-odhad). Podle Dhrymes (1994, odstavec 3.7) je

$$\sqrt{T} (\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}^0) = - \left[ \frac{1}{T} \frac{\partial^2 l^*}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}} (\bar{\boldsymbol{\delta}}) \right]^{-1} \frac{1}{\sqrt{T}} \frac{\partial l^*}{\partial \boldsymbol{\delta}} (\boldsymbol{\delta}^0), \quad (1.4.67)$$

$$\text{plim} \frac{1}{T} \frac{\partial^2 l^*}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}} (\bar{\boldsymbol{\delta}}) = -\mathbf{S}' (\tilde{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{Q}) \mathbf{S}, \quad (1.4.68)$$

$$\frac{1}{\sqrt{T}} \frac{\partial l^*}{\partial \boldsymbol{\delta}} (\boldsymbol{\delta}^0) \sim \mathbf{N}(\mathbf{0}, \mathbf{S}' (\tilde{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{Q}) \mathbf{S}), \quad (1.4.69)$$

kde  $\mathbf{Q}$  je konečná regulární matice ( $k \times k$ ) z předpokladu **(P3)**, resp. **(A3)**,  $\mathbf{S}$  je blokově diagonální matice s maticemi  $\mathbf{S}_1, \dots, \mathbf{S}_m$  na diagonále;  $\mathbf{S}_j = (\boldsymbol{\Pi}_j, \mathbf{L}_{2j})$  má rozměry ( $k \times (k_j + m_j - 1)$ ). Matice  $\mathbf{L}_{1j}$  typu ( $m \times (m_j - 1)$ ), resp.  $\mathbf{L}_{2j}$  typu ( $k \times k_j$ ),  $j = 1, \dots, m$ , obsahuje pouze 0 a 1 tak, aby platilo

$$\mathbf{Y}_j = \mathbf{Y} \mathbf{L}_{1j} [= (\mathbf{X} \boldsymbol{\Pi} + \mathbf{V}) \mathbf{L}_{1j} = \mathbf{X} \boldsymbol{\Pi}_j + \mathbf{V}_j], \quad \text{resp. } \mathbf{X}_j = \mathbf{X} \mathbf{L}_{2j}.$$

FIML-odhad je výpočetně mnohem náročnější než 3SLS-odhad, nicméně při existující apriorní informaci o parametrech soustavy je asymptoticky eficientní vůči 3SLS-odhadu, v opačném případě mají oba odhady shodnou asymptotickou rozptylovou matici. Lze rovněž ukázat, že FIML-odhad je konzistentní.

**Příklad 1.4.1.** Uvažujme plně rekurzivní soustavu simultánních rovnic (1.2.17), v ní jistě platí:

1.  $\boldsymbol{\Gamma}$  je horní trojúhelníková matice s hodnotami  $-1$  na diagonále, tj.  $|\det \boldsymbol{\Gamma}| = 1$  a  $\log |\det \boldsymbol{\Gamma}| = 0$ ,
2.  $\tilde{\boldsymbol{\Sigma}}$  je diagonální matice, takže  $\log (\det \tilde{\boldsymbol{\Sigma}}) = \sum_{j=1}^m \log \sigma_{jj}$ , dále

$$\text{tr} (\tilde{\boldsymbol{\Sigma}}^{-1} \mathbf{W}) = \sum_{j=1}^m \frac{1}{T \sigma_{jj}} (\mathbf{y}_j - \mathbf{Z}_j \boldsymbol{\delta}_j)' (\mathbf{y}_j - \mathbf{Z}_j \boldsymbol{\delta}_j).$$

Logaritmická věrohodnostní funkce (1.4.61) se zjevně zjednoduší na  $\log L = \sum_{j=1}^m \log L_j$ , kde

$$l_j = \log L_j = -\frac{T}{2} [\log(2\pi) + \log \sigma_{jj}] - \frac{1}{2\sigma_{jj}} (\mathbf{y}_j - \mathbf{Z}_j \boldsymbol{\delta}_j)' (\mathbf{y}_j - \mathbf{Z}_j \boldsymbol{\delta}_j).$$

Z předchozích vztahů snadno získáme

$$\frac{\partial l_j}{\partial \boldsymbol{\delta}_j} = \frac{1}{\sigma_{jj}} (\mathbf{Z}_j' \mathbf{Z}_j \boldsymbol{\delta}_j - \mathbf{Z}_j' \mathbf{y}_j) = \mathbf{0}, \quad (1.4.70)$$

$$\frac{\partial l_j}{\partial \sigma_{jj}} = -\frac{T}{2} \frac{1}{\sigma_{jj}} + \frac{1}{2\sigma_{jj}^2} (\mathbf{y}_j - \mathbf{Z}_j \boldsymbol{\delta}_j)' (\mathbf{y}_j - \mathbf{Z}_j \boldsymbol{\delta}_j) = 0. \quad (1.4.71)$$

Z (1.4.70) a (1.4.71) vidíme, že se FIML-odhad v rámci plně rekurzivní soustavy redukuje na OLS-odhad jednotlivých rovnic soustavy. Již dříve jsme uvedli,



že OLS-odhad je v tomto modelu konzistentní, nyní víme, že je rovněž asymptoticky eficientní.  $\triangle$

Jako druhý představíme *LIML-odhad* (*Limited Information Maximum Likelihood*), který, jak už jeho anglický název napovídá, plně nevyužívá dostupné informace z celé soustavy simultánních rovnic. Odvození logaritmické věrohodnostní funkce probíhá obdobně jako u FIML-odhadu, avšak nyní vycházíme pouze ze znalosti  $j$ -té rovnice soustavy, samozřejmě za předpokladu (asymptotické) normality její residuální složky. Seznámíme se zde jen se základními výsledky, podrobnější popis nalezneme v Dhrymes (1994, kapitola 4).

LIML-odhad parametru  $\delta_j$  v  $j$ -té rovnici soustavy

$$\mathbf{y}_j = \mathbf{Y}_j\boldsymbol{\gamma}_j + \mathbf{X}_j\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j = \mathbf{Z}_j\boldsymbol{\delta}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, m, \quad (1.4.72)$$

získáme řešením

$$\mathbf{Z}'_j(\mathbf{I}_T - \hat{\kappa}_j\mathbf{M}_X)(\mathbf{y}_j - \mathbf{Z}_j\boldsymbol{\delta}_j) = \mathbf{0}, \quad (1.4.73)$$

kde  $\mathbf{M}_X = \mathbf{I}_T - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$  je projekční matice, toto řešení můžeme explicitně vyjádřit

$$\hat{\boldsymbol{\delta}}_j = (\mathbf{Z}'_j(\mathbf{I}_T - \hat{\kappa}_j\mathbf{M}_X)\mathbf{Z}_j)^{-1} \mathbf{Z}'_j(\mathbf{I}_T - \hat{\kappa}_j\mathbf{M}_X)\mathbf{y}_j. \quad (1.4.74)$$

Konzistentně odhadnutá asymptotická rozptylová matice LIML-odhadu  $\hat{\boldsymbol{\delta}}_j$  je tvaru

$$\hat{\sigma}_{jj}^{LIML} (\mathbf{Z}'_j(\mathbf{I}_T - \hat{\kappa}_j\mathbf{M}_X)\mathbf{Z}_j)^{-1}, \quad (1.4.75)$$

$\hat{\sigma}_{jj}^{LIML}$  je konzistentním odhadem  $\sigma_{jj}$ ,

$$\hat{\sigma}_{jj}^{LIML} = \frac{(\mathbf{y}_j - \mathbf{Z}_j\hat{\boldsymbol{\delta}}_j)'(\mathbf{y}_j - \mathbf{Z}_j\hat{\boldsymbol{\delta}}_j)}{T}. \quad (1.4.76)$$

Konstantu  $\hat{\kappa}_j$  volíme tak, aby byl minimalizován podíl

$$\kappa_j = \frac{(\mathbf{y}_j - \mathbf{Y}_j\boldsymbol{\gamma}_j)' \mathbf{M}_{X_j} (\mathbf{y}_j - \mathbf{Y}_j\boldsymbol{\gamma}_j)}{(\mathbf{y}_j - \mathbf{Y}_j\boldsymbol{\gamma}_j)' \mathbf{M}_X (\mathbf{y}_j - \mathbf{Y}_j\boldsymbol{\gamma}_j)} = \frac{\boldsymbol{\alpha}'_j \mathbf{V}'_j \mathbf{M}_{X_j} \mathbf{V}_j \boldsymbol{\alpha}_j}{\boldsymbol{\alpha}'_j \mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j \boldsymbol{\alpha}_j}, \quad (1.4.77)$$

kde  $\boldsymbol{\alpha}_j = (-1, \boldsymbol{\gamma}'_j)'$ ,  $\mathbf{V}_j = (\mathbf{y}_j, \mathbf{Y}_j)$ ,  $\mathbf{M}_{X_j} = \mathbf{I}_T - \mathbf{X}_j(\mathbf{X}'_j\mathbf{X}_j)^{-1}\mathbf{X}'_j$  a  $\mathbf{M}_X = \mathbf{I}_T - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ . Poznamenejme, že  $\kappa_j \geq 1$ .

Vztah (1.4.77) zderivujeme podle parametru  $\boldsymbol{\alpha}_j$  a čitatele derivace následně položíme rovného  $\mathbf{0}$ , získáváme

$$2\mathbf{V}'_j \mathbf{M}_{X_j} \mathbf{V}_j \boldsymbol{\alpha}_j (\boldsymbol{\alpha}'_j \mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j \boldsymbol{\alpha}_j) - 2\mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j \boldsymbol{\alpha}_j (\boldsymbol{\alpha}'_j \mathbf{V}'_j \mathbf{M}_{X_j} \mathbf{V}_j \boldsymbol{\alpha}_j) = \mathbf{0}.$$

Obě strany předchozí rovnice vydělíme výrazem  $2\boldsymbol{\alpha}'_j \mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j \boldsymbol{\alpha}_j$ , tj.

$$\mathbf{V}'_j \mathbf{M}_{X_j} \mathbf{V}_j \boldsymbol{\alpha}_j - \kappa_j \mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j \boldsymbol{\alpha}_j = \mathbf{0}. \quad (1.4.78)$$

Dále celou rovnici zleva vynásobíme maticí  $(\mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j)^{-1/2}$ , tj.

$$((\mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j)^{-1/2} \mathbf{V}'_j \mathbf{M}_{X_j} \mathbf{V}_j (\mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j)^{-1/2} - \kappa_j \mathbf{I}_t) \boldsymbol{\alpha}_j^* = \mathbf{0}, \quad (1.4.79)$$

kde  $\boldsymbol{\alpha}_j^* = (\mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j)^{1/2} \boldsymbol{\alpha}_j$ .

Vidíme, že hledání  $\hat{\kappa}_j$ , minimální hodnoty podílu (1.4.77), se vztahem (1.4.79) převedlo na problematiku nalezení nejmenšího vlastního čísla reálné symetrické matice

$$(\mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j)^{-1/2} \mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j (\mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j)^{-1/2}. \quad (1.4.80)$$

Pro úplnost uvedme, že podle Wooldridge (2002, odstavec 12.5) je maximalizovaná hodnota logaritmické věrohodnostní funkce pro  $j$ -tou rovnici soustavy

$$-\frac{mT}{2} \log(2\pi) - \frac{T}{2} \log(\hat{\kappa}_j) - \frac{T}{2} \log |\mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j|. \quad (1.4.81)$$

**Poznámka 1.4.5.** LIML-odhad do jisté míry motivoval vznik *odhadu metodou nejmenšího podílu rozptylů* (*Least Variance Ratio, LVR*), který nevyžaduje apriorní znalost o typu rozdělení residuální složky  $j$ -té rovnice soustavy simultánních rovnic.

Konkrétně se snažíme nalézt odhad  $\mathbf{a}_j$  parametru  $\boldsymbol{\alpha}_j$  tak, aby  $\mathbf{V}_j \mathbf{a}_j$  bylo vysvětleno závislostí na  $\mathbf{X}_j$  téměř stejně dobře jako závislostí na  $\mathbf{X}$ . Naším úkolem je tedy zvolit  $\mathbf{a}_j$  tak, aby podíl residuálního rozptylu při závislosti  $\mathbf{V}_j \mathbf{a}_j$  na  $\mathbf{X}_j$  a residuálního rozptylu  $\mathbf{V}_j \mathbf{a}_j$  na  $\mathbf{X}$  byl minimální, tj. hledáme  $\mathbf{a}_j$  minimalizující podíl

$$\kappa_j = \frac{\mathbf{a}'_j \mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j \mathbf{a}_j}{\mathbf{a}'_j \mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j \mathbf{a}_j}. \quad (1.4.82)$$

Tuto úlohu jsme analyzovali v rámci hledání LIML-odhadu, viz výše. Požadovaný odhad  $\mathbf{a}_j$  získáme vyřešením rovnice (analogie (1.4.78))

$$(\mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j - \hat{\kappa}_j \mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j) \mathbf{a}_j = \mathbf{0}, \quad (1.4.83)$$

kde  $\hat{\kappa}_j$  je minimální hodnota výše uvedeného podílu.

LVR-odhad parametrů  $\boldsymbol{\gamma}_j$  a  $\boldsymbol{\beta}_j$  v  $j$ -té rovnici soustavy obdržíme v následujících krocích:

1. Spočteme  $\hat{\kappa}_j$  jakožto nejmenší vlastní číslo reálné symetrické matice

$$(\mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j)^{-1/2} \mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j (\mathbf{V}'_j \mathbf{M}_X \mathbf{V}_j)^{-1/2}.$$

2. Vyřešíme homogenní soustavu rovnic (1.4.83), nalezený vektor řešení znormujeme tak, aby první složkou byla  $-1$ , takto upravený vektor  $\mathbf{a}_j$  bereme za odhad parametru  $\boldsymbol{\alpha}_j = (-1, \boldsymbol{\gamma}'_j)'$ .
3. Odhad parametru  $\boldsymbol{\beta}_j$  nalezneme regreseí  $\mathbf{V}_j \mathbf{a}_j$  na  $\mathbf{X}_j$ .

Snadno nahlédneme, že LIML-odhad i LVR-odhad jsou totožné, ačkoliv jsme pro LVR-odhad nepožadovali apriorní znalost pravděpodobnostního rozdělení residuálních složek.  $\diamond$

Lze ukázat, že LIML-odhad (LVR-odhad) je konzistentní a má stejnou asymptotickou rozptylovou matici jako 2SLS-odhad, viz Dhrymes (1994, odstavec 4.7).

Na závěr odstavce poznamenejme, že zmiňované FIML-odhady i LIML-odhady jsou podle Wooldridge (2002, odstavec 12.5) invariantní vůči reparametrizaci modelu.

**Poznámka 1.4.6.** Dhrymes (1994, kapitola 4) ukázal, že  $\hat{\kappa}_j$ , minimální hodnota podílu (1.4.77), pro  $T \rightarrow \infty$  splňuje

$$\sqrt{T}(\hat{\kappa}_j - 1) \xrightarrow{d} 0,$$

navíc má  $T(\hat{\kappa}_j - 1)$  asymptoticky rozdělení  $\chi^2((k - k_j) - (m_j - 1))$ .

Na základě předchozích vztahů byl při konstrukci LIML-odhadu odvozen tzv. *Andersonův-Rubinův test*, ve kterém testujeme nulovou hypotézu

$$H_0 : h(\mathbf{\Pi}_{21}^j) \leq m_j - 1$$

proti alternativě

$$H_1 : h(\mathbf{\Pi}_{21}^j) = m_j,$$

příslušné značení bylo zavedeno v odstavci 1.2, srovnejme s (1.2.11). Asymptotický kritický obor testu je na hladině významnosti  $\alpha$  tvaru

$$T \cdot (\hat{\kappa}_j - 1) \geq \chi_{1-\alpha}^2((k - k_j) - (m_j - 1)). \quad (1.4.84)$$

Je zřejmé, že uvedený test úzce souvisí s problematikou identifikace, detailněji toto téma diskutuje Dhrymes (1994, odstavec 4.8).  $\diamond$

### 1.4.7 Odhady $k$ -té třídy

Odhady  $k$ -té třídy v sobě zahrnují celou skupinu odhadů, stručně se seznámíme pouze se základními principy a výsledky.

Odhad  $k$ -té třídy parametru  $\boldsymbol{\delta}_j$  v  $j$ -té rovnici soustavy

$$\mathbf{y}_j = \mathbf{Y}_j \boldsymbol{\gamma}_j + \mathbf{X}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j = \mathbf{Z}_j \boldsymbol{\delta}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, m, \quad (1.4.85)$$

je definován jako řešení rovnice

$$\mathbf{Z}'_j (\mathbf{I}_T - k \mathbf{M}_X) (\mathbf{y}_j - \mathbf{Z}_j \boldsymbol{\delta}_j) = \mathbf{0}, \quad (1.4.86)$$

kde  $\mathbf{M}_X = \mathbf{I}_T - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$  je projekční matice, toto můžeme explicitně vyjádřit

$$\hat{\boldsymbol{\delta}}_j = (\mathbf{Z}'_j (\mathbf{I}_T - k \mathbf{M}_X) \mathbf{Z}_j)^{-1} \mathbf{Z}'_j (\mathbf{I}_T - k \mathbf{M}_X) \mathbf{y}_j. \quad (1.4.87)$$

**Poznámka 1.4.7.** V předchozích částech textu jsme se již s některými reprezentanty odhadů  $k$ -té třídy setkali:

1. pro  $k = 0$  je  $\hat{\boldsymbol{\delta}}_j$  OLS-odhadem parametru  $\boldsymbol{\delta}_j$  v  $j$ -té rovnici soustavy,
2. pro  $k = \hat{\kappa}_j$  je  $\hat{\boldsymbol{\delta}}_j$  LVR-odhadem (za předpokladu normality residuální složky také LIML-odhadem) parametru  $\boldsymbol{\delta}_j$  v  $j$ -té rovnici soustavy,
3. pro  $k = 1$  je  $\hat{\boldsymbol{\delta}}_j$  2SLS-odhadem parametru  $\boldsymbol{\delta}_j$  v  $j$ -té rovnici soustavy.

$\diamond$

Lze ukázat, že při splnění obvyklých předpokladů jsou odhady  $k$ -té třídy konzistentní, jestliže  $\text{plim } k = 1$ , a konzistentní se stejnou asymptotickou rozptylovou maticí jako má 2SLS-odhad, jestliže  $\text{plim } \sqrt{T}(k - 1) = 0$ .

### 1.4.8 Shrnutí odhadových procedur

V předchozích odstavcích jsme se seznámili s několika nejrozšířenějšími metodami odhadu parametrů v rámci soustav simultánních rovnic. Obecně lze odhadové procedury v rámci uvažovaných soustav rozdělit do dvou skupin:

1. *Odhady s omezenou informací (Limited Information Method, LI)* odhadují jednotlivé rovnice soustavy samostatně, úplně nevyužijí dostupnou informaci. Mezi LI-odhady patří
  - ILS-odhad (odstavec 1.4.1),
  - IV-odhad (odstavec 1.4.2),
  - GMM-odhad pro jednotlivé rovnice soustavy (odstavec 1.4.3),
  - 2SLS-odhad (odstavec 1.4.4),
  - LIML-odhad (odstavec 1.4.6),
  - LVR-odhad (poznámka 1.4.5 v odstavci 1.4.6),
  - odhady  $k$ -té třídy (odstavec 1.4.7).
2. *Odhady s úplnou informací (Full Information Method, FI)* odhadují soustavu simultánních rovnic jako celek, úplně využijí veškerou dostupnou informaci, proto jsou asymptoticky eficientní vůči LI-odhadům, na druhé straně jsou výpočetně podstatně náročnější. Mezi FI-odhady patří
  - GMM-odhad pro celou soustavu rovnic (odstavec 1.4.3),
  - 3SLS-odhad (odstavec 1.4.5),
  - FIML-odhad (odstavec 1.4.6).

Všechny zde uvedené odhady jsou konzistentní a za poměrně obecných předpokladů asymptoticky normální (a asymptoticky nestranné).

V praxi se nejčastěji setkáme s 2SLS-odhadem a 3SLS-odhadem. 2SLS-odhad a LIML-odhad (LVR-odhad) mají shodnou asymptotickou rozptylovou matici, rovněž 3SLS-odhad a FIML-odhad mají stejnou asymptotickou rozptylovou matici.

Jsou-li k dispozici apriorní informace o rozptylové matici residuí  $\tilde{\Sigma}$ , je při využití této informace FIML-odhad asymptoticky eficientní vůči 3SLS-odhadu. FI-odhady jsou, jak již bylo řečeno, asymptoticky eficientní vůči LI-odhadům.

**Poznámka 1.4.8.** V případě, že je  $j$ -tá rovnice soustavy přesně identifikovaná, jsou ILS-odhad, 2SLS-odhad a LVR-odhad totožné, jestliže má navíc residuální složka normální rozdělení, shoduje se s nimi rovněž LIML-odhad.

Je-li rozptylová matice residuí  $\tilde{\Sigma}$  diagonální nebo jsou-li všechny rovnice soustavy přesně identifikované, jsou 2SLS-odhad a 3SLS-odhad stejné.  $\diamond$

**Poznámka 1.4.9.** Ve všech odhadových procedurách prezentovaných v předchozím odstavci 1.4 jsme při výpočtu konzistentních odhadů parametru  $\sigma_{ij}$  dělili číslem  $T$ . Tento postup bývá aplikován v dostupném softwaru, navíc je běžně zmiňován v literatuře. Je však vhodné uvést, že namísto čísla  $T$  je v rámci (konzistentního) odhadu parametru  $\sigma_{ij}$  možné dělit například těmito čísly

$$T - \max \{k_i + m_i - 1, k_j + m_j - 1\} \quad \text{nebo}$$

$$\sqrt{(T - (k_i + m_i - 1)) \cdot (T - (k_j + m_j - 1))}.$$

◇

Uvedené vlastnosti odhadů se však zpravidla projevují pouze asymptoticky. V případě malého počtu pozorování se jejich vlastnosti vyšetřují na základě simulací (metoda *Monte Carlo*). Simulační metody jsou v moderní ekonometrii velmi populární, často se také používá statistická metoda *bootstrap*, která se s nedostatkem pozorování vypořádává tím, že generuje data z rozdělení odhadnutého na základě původních dat.

**Poznámka 1.4.10.** Zajímá-li nás odhad matice parametrů redukovaného tvaru soustavy simultánních rovnic (1.1.7)  $\Pi = -B\Gamma^{-1}$ , můžeme jej na základě odstavce 1.4.1 zkonstruovat jako konzistentní OLS-odhad  $P = (X'X)^{-1}X'Y$ . Pokud však máme k dispozici  $\hat{B}$  a  $\hat{\Gamma}$  konzistentní odhady parametrů  $B$  a  $\Gamma$ , je podstatně výhodnější použít tzv. *odvozený odhad parametrů redukovaného tvaru*

$$\hat{\Pi} = -\hat{B}\hat{\Gamma}^{-1}, \quad (1.4.88)$$

který je konzistentní a relativně eficientní vůči odhadu  $P$ , tj.  $P - \hat{\Pi}$  je pozitivně semidefinitní matice, a to z důvodu využití apriorních informací o strukturálních parametrech. Více viz Cipra (1984, odstavec 13.1). ◇

## 1.5 Testy exogenity

V praktických aplikacích se někdy setkáváme s otázkou, které proměnné  $j$ -té rovnice soustavy simultánních rovnic můžeme považovat (na základě teoretických poznatků z ekonomie, financí, apod.) za exogenní. K tomuto účelu se využívají tzv. *testy exogenity*, s nimiž se v tomto odstavci seznámíme. Doplňující informace nalezneme například v Greene (2003, odstavec 15.8) nebo v Cipra (2008, odstavec 7.4.2).

Nejpoužívanějším testem exogenity je *Hausmanův test*, který ověřuje nulovou hypotézu, zda proměnné  $(Y_j, X_j)$  v  $j$ -té rovnici

$$y_j = Y_j\gamma_j + X_j\beta_j + \varepsilon_j, \quad j = 1, \dots, m, \quad (1.5.1)$$

mohou být považovány za exogenní, proti alternativní hypotéze, že proměnné  $Y_j$  nejsou exogenní. K tomu nám stačí v rovnici

$$y_j = \hat{Y}_j\gamma_j + X_j\beta_j + \varepsilon_j, \quad \text{kde } \hat{Y}_j = X_j(X_j'X_j)^{-1}X_j'Y_j, \quad (1.5.2)$$

provést test nulové hypotézy  $H_0 : \boldsymbol{\gamma}_j = \mathbf{0}$  proti alternativě  $H_1 : \boldsymbol{\gamma}_j \neq \mathbf{0}$ .

K ověření uvedené hypotézy můžeme za předpokladu (asymptotické) normality pro  $m_j > 1$  použít:

1.  $F$ -test s kritickým oborem na hladině významnosti  $\alpha$

$$\frac{T - (k_j + m_j - 1)}{m_j - 1} \cdot \frac{RRSS_j - URSS_j}{URSS_j} \geq F_{1-\alpha}(m_j - 1, T - (k_j + m_j - 1)), \quad (1.5.3)$$

kde  $URSS_j$  je residuální součet čtverců v modelu (1.5.2) a  $RRSS_j$  je residuální součet čtverců v tomtéž modelu, ve kterém jsme aplikovali  $\boldsymbol{\gamma}_j = \mathbf{0}$ .

2.  $LM$ -test (viz Cipra (2008, odstavec 5.4)) s (asymptotickým) kritickým oborem na hladině významnosti  $\alpha$

$$LM_j = T \cdot R_j^2 \geq \chi_{1-\alpha}^2(m_j - 1), \quad (1.5.4)$$

kde  $R_j^2$  je koeficient determinace v modelu

$$\hat{\boldsymbol{\varepsilon}}_j = \hat{\mathbf{Y}}_j \boldsymbol{\gamma}_j + \mathbf{X}_j \boldsymbol{\beta}_j + \mathbf{u}_j, \quad (1.5.5)$$

pro vypočtená OLS-residua  $\hat{\boldsymbol{\varepsilon}}_j$  v  $\mathbf{y}_j = \mathbf{X}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j$ .

## 1.6 Testy specifikace

V rámci soustav simultánních rovnic můžeme rovněž diskutovat problém spočívající v tom, že se v daném modelu mohou objevit proměnné, které považujeme za exogenní, avšak tyto jsou současně korelovány s některými residuálními složkami, zjevně tak porušují předpoklady uvažované v úvodu této kapitoly. V uvedeném kontextu pak zpravidla hovoříme o *problému specifikace modelu*.

K odhalení zmíněné situace slouží např. *Hausmanův test specifikace*, který ověřuje nulovou hypotézu, že jsou exogenní proměnné v jednotlivých rovnicích soustavy současně nekorelované s residuálními složkami všech rovnic modelu, proti alternativě, že tomu tak není. Za platnosti nulové hypotézy jsou  $\hat{\boldsymbol{\delta}}$  2SLS-odhad a  $\hat{\boldsymbol{\delta}}^*$  3SLS-odhad parametru  $\boldsymbol{\delta}$  konzistentní,  $\hat{\boldsymbol{\delta}}^*$  je navíc asymptoticky eficientní. V opačném případě je aspoň jeden z odhadů  $\hat{\boldsymbol{\delta}}$ ,  $\hat{\boldsymbol{\delta}}^*$  nekonzistentní.

Za platnosti nulové hypotézy obdržíme Waldovu statistiku (viz Cipra (2008, odstavec 5.4))

$$W = (\hat{\boldsymbol{\delta}}^* - \hat{\boldsymbol{\delta}})' \left[ \widehat{\text{var}}(\hat{\boldsymbol{\delta}}) - \widehat{\text{var}}(\hat{\boldsymbol{\delta}}^*) \right]^{-1} (\hat{\boldsymbol{\delta}}^* - \hat{\boldsymbol{\delta}}), \quad (1.6.1)$$

která má asymptoticky  $\chi^2(p)$  rozdělení;  $\widehat{\text{var}}(\hat{\boldsymbol{\delta}})$  a  $\widehat{\text{var}}(\hat{\boldsymbol{\delta}}^*)$  jsou konzistentní odhady příslušných rozptylových matic, viz (1.4.48) - (1.4.49) a (1.4.56),  $\left[ \widehat{\text{var}}(\hat{\boldsymbol{\delta}}) - \widehat{\text{var}}(\hat{\boldsymbol{\delta}}^*) \right]^{-1}$  je zobecněná pseudoinverzní matice k matici  $\left[ \widehat{\text{var}}(\hat{\boldsymbol{\delta}}) - \widehat{\text{var}}(\hat{\boldsymbol{\delta}}^*) \right]$  a počet stupňů volnosti  $p$  je roven hodnotě uvedené zobecněné pseudoinverzní matice.

Připomeňme, že *zobecněnou pseudoinverzní maticí* (tzv. *Mooreovou-Penroseovou pseudoinverzí*) k matici  $\mathbf{A}$  typu  $(m \times n)$  rozumíme (jednoznačně) určenou matici  $\mathbf{A}^-$  typu  $(n \times m)$ , pro kterou platí

$$\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}, \quad \mathbf{A}^-\mathbf{A}\mathbf{A}^- = \mathbf{A}^-, \quad \mathbf{A}^-\mathbf{A} \text{ a } \mathbf{A}\mathbf{A}^- \text{ jsou symetrické.}$$

**Poznámka 1.6.1.** V případě právě popsaného Hausmanova testu se zpravidla obtížně určuje, které exogenní proměnné nastíněný problém specifikace modelu způsobují, z tohoto důvodu byl navržen *Spencerův-Berkův test specifikace* pro jednotlivé rovnice soustavy. Mějme  $j$ -tou rovnici soustavy

$$\mathbf{y}_j = \mathbf{Y}_j \boldsymbol{\gamma}_j + \mathbf{X}_j \boldsymbol{\beta}_j + \mathbf{x}^e \theta_j + \boldsymbol{\varepsilon}_j = (\mathbf{Y}_j, \mathbf{X}_j, \mathbf{x}^e) \boldsymbol{\delta}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, m, \quad (1.6.2)$$

kde  $\mathbf{x}^e$  je proměnná, o které se snažíme zjistit, zda není příčinou problému specifikace modelu. Ověřujeme nulovou hypotézu

$H_0$  : „ $\hat{\boldsymbol{\delta}}_j$  a  $\hat{\boldsymbol{\delta}}_j^*$  jsou konzistentní a  $\hat{\boldsymbol{\delta}}_j^*$  je asymptoticky eficientní“ ( $\mathbf{x}^e$  je exogenní) proti alternativě

$H_1$  : „ $\hat{\boldsymbol{\delta}}_j$  je konzistentní,  $\hat{\boldsymbol{\delta}}_j^*$  je nekonzistentní“ ( $\mathbf{x}^e$  je endogenní).

Za platnosti nulové hypotézy uvažujeme Waldovu statistiku

$$W^* = (\hat{\boldsymbol{\delta}}_j^* - \hat{\boldsymbol{\delta}}_j)' \left[ \widehat{\text{var}}(\hat{\boldsymbol{\delta}}_j) - \widehat{\text{var}}(\hat{\boldsymbol{\delta}}_j^*) \right]^{-1} (\hat{\boldsymbol{\delta}}_j^* - \hat{\boldsymbol{\delta}}_j), \quad (1.6.3)$$

která má asymptoticky  $\chi^2(1)$  rozdělení;  $\widehat{\text{var}}(\hat{\boldsymbol{\delta}}_j)$  a  $\widehat{\text{var}}(\hat{\boldsymbol{\delta}}_j^*)$  jsou konzistentní odhady příslušných rozptylových matic,  $\left[ \widehat{\text{var}}(\hat{\boldsymbol{\delta}}_j) - \widehat{\text{var}}(\hat{\boldsymbol{\delta}}_j^*) \right]^{-1}$  je zobecněná pseudoinverzní matice k matici  $\left[ \widehat{\text{var}}(\hat{\boldsymbol{\delta}}_j) - \widehat{\text{var}}(\hat{\boldsymbol{\delta}}_j^*) \right]$ .

$\hat{\boldsymbol{\delta}}_j^*$  je 2SLS-odhad parametru  $\boldsymbol{\delta}_j$  v rámci (1.6.2), přičemž se  $\mathbf{x}^e$  považuje za exogenní proměnnou.  $\hat{\boldsymbol{\delta}}_j$  je IV-odhad parametru  $\boldsymbol{\delta}_j$  založený na regresi  $\mathbf{y}_j$  na  $\mathbf{Y}_j$ ,  $\mathbf{X}_j$  a  $\hat{\mathbf{x}}^e$ , kde  $\hat{\mathbf{x}}^e$  jsou vypočtené OLS-hodnoty v regresi proměnné  $\mathbf{x}^e$  na zbylé exogenní proměnné.

Uvedený Spencerův-Berkův test lze zřejmým způsobem rozšířit i pro sady proměnných, které by mohly v  $j$ -té rovnici soustavy způsobovat problém specifikace.

Další podrobnosti k testům specifikace modelu lze nalézt v Greene (2003, odstavce 5.5 a 15.8), popř. v Dhrymes (1994, kapitola 1).  $\diamond$

## 1.7 Dynamické soustavy simultánních rovnic

*Dynamickou soustavou simultánních rovnic* rozumíme soustavu

$$y_{jt} = \sum_{i=1, i \neq j}^m \gamma_{ji} y_{it} + \sum_{i=1}^k \beta_{ji} x_{it} + \varepsilon_{jt}, \quad j = 1, \dots, m, \quad t = 1, \dots, T, \quad (1.7.1)$$

ve které jsou na pravé straně  $j$ -té rovnice kromě endogenních proměnných  $y_{it}$  ( $i \neq j$ ) a exogenních proměnných  $x_{it}$  obsaženy rovněž zpožděné endogenní proměnné typu  $y_{i,t-k}$  ( $k \geq 1$ ). Tyto zpožděné proměnné jsou současně nekorelované s residuální složkou, tudíž je lze v rámci uvažované rovnice formálně zařadit do skupiny exogenních proměnných. Takto vzniklé exogenní proměnné označujeme jako *predeterminované*, protože vznikly v rámci uvažované soustavy v minulých časech, zbylé

exogenní proměnné, které vznikly naprosto vně uvažované soustavy, nazýváme jako *striktně exogenní*, viz strana 8.

Vzhledem k předchozí diskuzi se v rámci dynamických soustav simultánních rovnic v nezměněné podobě aplikují odhadové postupy uvedené v odstavci 1.4.

*Strukturální tvar dynamické soustavy simultánních rovnic* můžeme zapsat jako

$$\mathbf{y}_t \mathbf{\Gamma} + \sum_{k=1}^p L^k(\mathbf{y}_t) \mathbf{\Phi}_k + \mathbf{x}_t \mathbf{B} + \boldsymbol{\varepsilon}_t = \mathbf{0}, \quad t = p+1, \dots, T, \quad (1.7.2)$$

kde  $p \in \{1, \dots, T-1\}$ ,  $L$  označuje lineární operátor zpětného časového posunutí, tj. pro  $k \in \mathbb{N}_0$  platí  $L^k(\mathbf{y}_t) = \mathbf{y}_{(t-k)\cdot}$ ,  $\mathbf{\Phi}_k$  je matice  $(m \times m)$  parametrů a  $\mathbf{x}_t$  jsou striktně exogenní proměnné.

*Redukovaný tvar soustavy simultánních rovnic* získáme vynásobením vztahu (1.7.2) zprava maticí  $\mathbf{\Gamma}^{-1}$ , tj.

$$\mathbf{y}_t = \mathbf{x}_t \mathbf{\Pi} + \sum_{k=1}^p L^k(\mathbf{y}_t) \mathbf{\Delta}_k + \mathbf{v}_t, \quad (1.7.3)$$

kde  $\mathbf{\Pi} = -\mathbf{B}\mathbf{\Gamma}^{-1}$ ,  $\mathbf{\Delta}_k = -\mathbf{\Phi}_k\mathbf{\Gamma}^{-1}$  a  $\mathbf{v}_t = -\boldsymbol{\varepsilon}_t\mathbf{\Gamma}^{-1}$ , existuje-li inverzní matice k  $\mathbf{\Gamma}$ .

Vztah (1.7.2) lze přidáním identit tvaru  $\mathbf{y}_{(t-1)\cdot} = \mathbf{y}_{(t-1)\cdot}, \dots, \mathbf{y}_{(t-p+1)\cdot} = \mathbf{y}_{(t-p+1)\cdot}$  pro  $p > 1$  jednodušeji vyjádřit jako

$$\mathbf{y}_t^* \mathbf{\Gamma}^* + \mathbf{x}_t \mathbf{B}^* + \mathbf{y}_{(t-1)\cdot}^* \mathbf{\Phi}^* + \boldsymbol{\varepsilon}_t^* = \mathbf{0}, \quad t = p+1, \dots, T, \quad (1.7.4)$$

kde platí

$$\mathbf{y}_t^* = (\mathbf{y}_t, \dots, \mathbf{y}_{(t-p+1)\cdot}), \quad \mathbf{\Gamma}^* = \begin{pmatrix} \mathbf{\Gamma} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_m & \dots & \mathbf{0} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I}_m \end{pmatrix}, \quad \mathbf{B}^* = (\mathbf{B}, \mathbf{0}, \dots, \mathbf{0}),$$

$$\mathbf{\Phi}^* = \begin{pmatrix} \mathbf{\Phi}_1 & -\mathbf{I}_m & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{\Phi}_2 & \mathbf{0} & -\mathbf{I}_m & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \mathbf{\Phi}_{p-1} & \mathbf{0} & \mathbf{0} & \dots & -\mathbf{I}_m \\ \mathbf{\Phi}_p & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix}, \quad \boldsymbol{\varepsilon}_t^* = (\boldsymbol{\varepsilon}_t, \mathbf{0}, \dots, \mathbf{0}).$$

Pro  $p = 1$  v uvažované konstrukci (1.7.4) zjevně pokládáme

$$\mathbf{y}_t^* = \mathbf{y}_t, \quad \mathbf{\Gamma}^* = \mathbf{\Gamma}, \quad \mathbf{B}^* = \mathbf{B}, \quad \mathbf{\Phi}^* = \mathbf{\Phi} \quad \text{a} \quad \boldsymbol{\varepsilon}_t^* = \boldsymbol{\varepsilon}_t.$$

Podobně jako výše je možné z (1.7.4) obdržet redukovaný tvar soustavy

$$\mathbf{y}_t^* = \mathbf{x}_t \mathbf{\Pi}^* + \mathbf{y}_{(t-1)\cdot}^* \mathbf{\Delta}^* + \mathbf{v}_t^*, \quad (1.7.5)$$

kde  $\mathbf{\Pi}^* = -\mathbf{B}^*(\mathbf{\Gamma}^*)^{-1}$ ,  $\mathbf{\Delta}^* = -\mathbf{\Phi}^*(\mathbf{\Gamma}^*)^{-1}$  a  $\mathbf{v}_t^* = -\boldsymbol{\varepsilon}_t^*(\mathbf{\Gamma}^*)^{-1}$ , existuje-li inverzní matice k  $\mathbf{\Gamma}$ .



Dosadíme-li za  $\mathbf{y}_{(t-1)}^*$  v (1.7.5) týž vztah posunutý v čase o jednotku zpět, dostáváme

$$\mathbf{y}_t^* = \mathbf{x}_t \cdot \mathbf{\Pi}^* + \mathbf{x}_{(t-1)} \cdot \mathbf{\Pi}^* \Delta^* + \mathbf{y}_{(t-2)}^* \cdot \Delta^{*2} + \mathbf{v}_t^* + \mathbf{v}_{(t-1)}^* \cdot \Delta^*.$$

Opakováním tohoto postupu pro přípustná  $t$  získáme

$$\mathbf{y}_t^* = \sum_{s=0}^{t-p-1} \mathbf{x}_{(t-s)} \cdot \mathbf{\Pi}^* \Delta^{*s} + \mathbf{y}_p^* \cdot \Delta^{*t-p} + \sum_{s=0}^{t-p-1} \mathbf{v}_{(t-s)}^* \cdot \Delta^{*s}. \quad (1.7.6)$$

Odtud snadno nahlédneme, jakým způsobem proměnné  $\mathbf{y}_p^*$ ,  $\mathbf{x}_{(t-s)}$  a residuální složky  $\mathbf{v}_{(t-s)}^*$  určují současnou hodnotu endogenních proměnných  $\mathbf{y}_t^*$ .

**Poznámka 1.7.1.** V rámci uvedeného modelu (1.7.6) lze rozlišit:

- *okamžitý vliv (short-run effect)* změny vysvětlujících proměnných, který je určen jako

$$\frac{\partial y_{et}^*}{\partial x_{ft}} = (\mathbf{\Pi}^*)_{fe}, \quad (1.7.7)$$

$\mathbf{\Pi}^*$  je matice tzv. *impact (short-run) multipliers*;

- *kumulovaný vliv* změny vysvětlujících proměnných až do zpoždění  $\tau$ , který je určen součtem

$$\sum_{s=0}^{\tau} \mathbf{\Pi}^* \Delta^{*s}, \quad (1.7.8)$$

$\mathbf{\Pi}^* \Delta^{*s}$  je matice tzv. *dynamic multipliers*, neboť

$$\frac{\partial y_{et}^*}{\partial x_{f,t-s}} = (\mathbf{\Pi}^* \Delta^{*s})_{fe}; \quad (1.7.9)$$

- *dlouhodobý vliv (long-run effect)* změny vysvětlujících proměnných, který je dán nekonečným součtem

$$\sum_{s=0}^{\infty} \mathbf{\Pi}^* \Delta^{*s}, \quad (1.7.10)$$

vzniklá matice bývá označována jako matice tzv. *equilibrium (long-run) multipliers*.

Navíc, předpokládáme-li  $\lim_{t \rightarrow \infty} \Delta^{*t} = \mathbf{0}$  ve smyslu konvergence jednotlivých prvků matice  $\Delta^*$ , dostáváme

$$\sum_{s=0}^{\infty} \mathbf{\Pi}^* \Delta^{*s} = \mathbf{\Pi}^* [\mathbf{I}_m + \Delta^* + \Delta^{*2} + \dots] = \mathbf{\Pi}^* [\mathbf{I}_m - \Delta^*]^{-1}. \quad (1.7.11)$$

Kumulovaný vliv až do zpoždění  $\tau$  pak lze jednodušeji vyjádřit jako

$$\sum_{s=0}^{\tau} \mathbf{\Pi}^* \Delta^{*s} = \mathbf{\Pi}^* [\mathbf{I}_m - \Delta^*]^{-1} [\mathbf{I}_m - \Delta^{*\tau}]. \quad (1.7.12)$$

◇

V poznámce 1.7.1 jsme požadovali  $\lim_{t \rightarrow \infty} \Delta^{*t} = \mathbf{0}$ , podívejme se nyní na tento předpoklad podrobněji.

Uvažujeme spektrální rozklad matice  $\Delta^* = -\Phi^*(\Gamma^*)^{-1}$ ,

$$\Delta^* = C\Lambda C', \quad (1.7.13)$$

kde  $\Lambda$  je diagonální matice vlastních čísel matice  $\Delta^*$ ,  $C$  je ortogonální matice, tj.  $C'C = CC' = I_m$ , jejíž sloupce tvoří ortonormální vlastní vektory příslušné vlastním číslům matice  $\Delta^*$ .

Snadno nahlédneme

$$\Delta^{*t} = C\Lambda^t C',$$

odtud je zřejmé, že požadavek  $\lim_{t \rightarrow \infty} \Delta^{*t} = \mathbf{0}$  je ekvivalentní  $\lim_{t \rightarrow \infty} \Lambda^t = \mathbf{0}$ , což je splněno, jestliže pro největší vlastní číslo  $\lambda_{max} \in \mathbb{C}$  matice  $\Delta^*$  platí  $|\lambda_{max}| < 1$ .

Tuto podmínku nazýváme *podmínkou stability* pro dynamické soustavy simultánních rovnic.

**Poznámka 1.7.2.** Dynamické soustavy simultánních rovnic lze také chápat jako speciální případ modelů VAR (*vektorová autoregrese*), podrobněji se této problematice věnuje Cipra (2008, kapitola 12). ◇

## 1.8 Předpovědi v rámci soustav simultánních rovnic

V tomto odstavci se budeme zabývat konstrukcí lineárních předpovědí v rámci ekonometrických soustav simultánních rovnic, z aplikačního hlediska se zjevně jedná o velmi důležitou dovednost.

Z přirozených důvodů je možné předpovídat pouze z redukováného tvaru soustavy (nedynamické či dynamické). V praxi lze předpovědní proceduru shrnout do čtyř kroků:

1. Odhad parametrů redukováného tvaru soustavy (viz poznámka 1.4.10).
2. Volba předpovědního horizontu.
3. Volba hodnot exogenních proměnných podle předpovědního horizontu.
4. Vlastní výpočet předpovědi.

### 1.8.1 Nedynamické modely

V rámci redukováného tvaru (nedynamické) soustavy simultánních rovnic (1.1.7)

$$\mathbf{y}_t = \mathbf{x}_t \mathbf{\Pi} + \mathbf{v}_t, \quad t = 1, \dots, T, \quad (1.8.1)$$

obdržíme předpověď  $\hat{\mathbf{y}}_{(T+\tau)\cdot}(T)$  z času  $T$  o  $\tau \in \mathbb{N}$  kroků dopředu jako

$$\hat{\mathbf{y}}_{(T+\tau)\cdot}(T) = \mathbf{x}_{(t+\tau)\cdot} \hat{\mathbf{\Pi}}, \quad (1.8.2)$$

kde  $\hat{\mathbf{\Pi}}$  značí odhad matice parametrů  $\mathbf{\Pi}$  a kde, jak již bylo naznačeno v úvodu tohoto odstavce, předpokládáme znalost hodnot striktně exogenních proměnných  $\mathbf{x}_t$ ,  $t = 1, \dots, T + \tau$ .

Chybou předpovědi (1.8.2) rozumíme

$$\hat{\mathbf{e}}_{(T+\tau)\cdot}(T) = \mathbf{y}_{(T+\tau)\cdot} - \hat{\mathbf{y}}_{(T+\tau)\cdot}(T) = \mathbf{v}_{(T+\tau)\cdot} - \mathbf{x}_{(T+\tau)\cdot} (\hat{\mathbf{\Pi}} - \mathbf{\Pi}). \quad (1.8.3)$$

Položíme-li  $\boldsymbol{\pi} = (\pi'_{\cdot 1}, \dots, \pi'_{\cdot m})'$  a  $\hat{\boldsymbol{\pi}} = (\hat{\pi}'_{\cdot 1}, \dots, \hat{\pi}'_{\cdot m})'$ , kde  $\pi_{\cdot j}$  a  $\hat{\pi}_{\cdot j}$  označují  $j$ -té sloupce matic  $\mathbf{\Pi}$  a  $\hat{\mathbf{\Pi}}$ , můžeme poslední člen předcházejícího vztahu zapsat jako

$$\mathbf{x}_{(T+\tau)\cdot} (\hat{\mathbf{\Pi}} - \mathbf{\Pi}) = [(\mathbf{I}_m \otimes \mathbf{x}_{(T+\tau)\cdot}) \cdot (\hat{\boldsymbol{\pi}} - \boldsymbol{\pi})]'$$

Jestliže je  $\hat{\mathbf{\Pi}}$  nestranným odhadem matice parametrů  $\mathbf{\Pi}$ , potom platí

$$\mathbf{E}(\hat{\mathbf{e}}'_{(T+\tau)\cdot}(T)) = \mathbf{0}, \quad (1.8.4)$$

$$\text{var}(\hat{\mathbf{e}}'_{(T+\tau)\cdot}(T)) = \tilde{\boldsymbol{\Omega}} + (\mathbf{I}_m \otimes \mathbf{x}_{(T+\tau)\cdot}) \text{var}(\hat{\boldsymbol{\pi}}) (\mathbf{I}_m \otimes \mathbf{x}_{(T+\tau)\cdot})'. \quad (1.8.5)$$

Uvažujme  $\hat{\mathbf{e}}_{(T+\tau)\cdot}^{(i)}(T)$ ,  $i = 1, 2$ , předpovědní chybu založenou na nestranném odhadu  $\hat{\mathbf{\Pi}}^{(i)}$  matice parametrů  $\mathbf{\Pi}$ . Porovnejme nyní rozptylové matice předpovědních chyb  $\boldsymbol{\Sigma}_{\tau i} = \text{var}(\hat{\mathbf{e}}_{(T+\tau)\cdot}^{(i)}(T))$ ,  $i = 1, 2$ ,

$$\boldsymbol{\Sigma}_{\tau 2} - \boldsymbol{\Sigma}_{\tau 1} = (\mathbf{I}_m \otimes \mathbf{x}_{(T+\tau)\cdot}) \left[ \text{var}(\hat{\boldsymbol{\pi}}^{(2)}) - \text{var}(\hat{\boldsymbol{\pi}}^{(1)}) \right] (\mathbf{I}_m \otimes \mathbf{x}_{(T+\tau)\cdot})'. \quad (1.8.6)$$

Poznamenejme, že předpověď založená na  $\hat{\mathbf{\Pi}}^{(2)}$  je *relativně eficientní* vůči předpovědi založené na  $\hat{\mathbf{\Pi}}^{(1)}$ , jestliže  $\boldsymbol{\Sigma}_{\tau 2} - \boldsymbol{\Sigma}_{\tau 1} \geq 0$ , tj. jedná se o pozitivně semidefinitní matici.

Ze vztahu (1.8.6) vyplývá, že předpověď založená na  $\hat{\mathbf{\Pi}}^{(2)}$  je relativně eficientní vůči předpovědi založené na  $\hat{\mathbf{\Pi}}^{(1)}$ , právě když je odhad  $\hat{\mathbf{\Pi}}^{(2)}$  relativně eficientní vůči odhadu  $\hat{\mathbf{\Pi}}^{(1)}$ . Je rovněž zřejmé, že relativní efience předpovědí nezávisí na předpovědním horizontu  $\tau$ .

## 1.8.2 Dynamické modely

Mějme redukovaný tvar dynamické soustavy simultánních rovnic (1.7.5)

$$\mathbf{y}_t^* = \mathbf{x}_t \mathbf{\Pi}^* + \mathbf{y}_{(t-1)}^* \boldsymbol{\Delta}^* + \mathbf{v}_t^*, \quad t = p + 1, \dots, T, \quad (1.8.7)$$

za  $\mathbf{y}_{(t-1)}^*$  postupně dosazujeme uvedený zpožděný vztah, pro přípustná  $\tau \in \mathbb{N}$  dostáváme

$$\mathbf{y}_t^* = \mathbf{y}_{(t-\tau)}^* \boldsymbol{\Delta}^{*\tau} + \sum_{i=0}^{\tau-1} \mathbf{x}_{(t-i)} \mathbf{\Pi}^* \boldsymbol{\Delta}^{*i} + \sum_{i=0}^{\tau-1} \mathbf{v}_{(t-i)}^* \boldsymbol{\Delta}^{*i}. \quad (1.8.8)$$

Rovnost (1.8.8) vynásobíme zprava maticí  $\mathbf{J} = (\mathbf{I}_m, \mathbf{0}, \dots, \mathbf{0})'$  typu  $(mp \times m)$  ( $p$  je definováno v odstavci 1.7) a namísto  $t$  položíme  $T + \tau$ , obdržíme

$$[\mathbf{y}'_{(T+\tau)} \mathbf{J} =] \mathbf{y}_{(T+\tau)} = \mathbf{y}_T^* \Delta^{*\tau} \mathbf{J} + \sum_{i=0}^{\tau-1} \mathbf{x}_{(T+\tau-i)} \Pi^* \Delta^{*i} \mathbf{J} + \sum_{i=0}^{\tau-1} \mathbf{v}_{(T+\tau-i)} \mathbf{J}' \Delta^{*i} \mathbf{J}, \quad (1.8.9)$$

neboť platí  $\mathbf{v}_t^* = \mathbf{v}_t \mathbf{J}'$ .

Nejlepší lineární předpověď (ve smyslu relativní eficeince *střední čtvercové chyby MSE*) je při znalosti hodnot striktně exogenních proměnných  $\mathbf{x}_t$ ,  $t = p+1, \dots, T+\tau$ , podle Lütkepohl (2005, kapitola 2 a 10) tvaru

$$\mathbf{y}_{(T+\tau)}(T) = \mathbf{y}_T^* \Delta^{*\tau} \mathbf{J} + \sum_{i=0}^{\tau-1} \mathbf{x}_{(T+\tau-i)} \Pi^* \Delta^{*i} \mathbf{J}. \quad (1.8.10)$$

Chybou předpovědi je

$$\mathbf{e}_{(T+\tau)}(T) = \mathbf{y}_{(T+\tau)} - \mathbf{y}_{(T+\tau)}(T) = \sum_{i=0}^{\tau-1} \mathbf{v}_{(T+\tau-i)} \mathbf{J}' \Delta^{*i} \mathbf{J}, \quad (1.8.11)$$

přičemž snadno nahlédneme

$$\mathbf{E}(\mathbf{e}'_{(T+\tau)}(T)) = \mathbf{0}, \quad (1.8.12)$$

$$\text{MSE}(\mathbf{e}'_{(T+\tau)}(T)) = \mathbf{E}(\mathbf{e}'_{(T+\tau)}(T) \mathbf{e}_{(T+\tau)}(T)) = \sum_{i=0}^{\tau-1} \mathbf{J}' \Delta^{*i} \mathbf{J} \tilde{\Omega} \mathbf{J}' (\Delta^{*i})' \mathbf{J}. \quad (1.8.13)$$

Předpovědi na základě odhadů  $\hat{\Pi}^*$ ,  $\hat{\Delta}^*$  parametrů  $\Pi^*$ ,  $\Delta^*$  chápeme

$$\hat{\mathbf{y}}_{(T+\tau)}(T) = \mathbf{y}_T^* \hat{\Delta}^{*\tau} \mathbf{J} + \sum_{i=0}^{\tau-1} \mathbf{x}_{(T+\tau-i)} \hat{\Pi}^* \hat{\Delta}^{*i} \mathbf{J}. \quad (1.8.14)$$

Předpovědní chybu pak můžeme vyjádřit jako

$$\hat{\mathbf{e}}_{(T+\tau)}(T) = [\mathbf{y}_{(T+\tau)} - \mathbf{y}_{(T+\tau)}(T)] + [\mathbf{y}_{(T+\tau)}(T) - \hat{\mathbf{y}}_{(T+\tau)}(T)], \quad (1.8.15)$$

$$\begin{aligned} \text{MSE}(\hat{\mathbf{e}}'_{(T+\tau)}(T)) &= \mathbf{E}(\hat{\mathbf{e}}'_{(T+\tau)}(T) \hat{\mathbf{e}}_{(T+\tau)}(T)) = \text{MSE}(\mathbf{e}'_{(T+\tau)}(T)) + \\ &+ \text{MSE} \left[ (\mathbf{y}_{(T+\tau)}(T) - \hat{\mathbf{y}}_{(T+\tau)}(T))' \right], \end{aligned} \quad (1.8.16)$$

neboť  $\mathbf{y}_{(T+\tau)} - \mathbf{y}_{(T+\tau)}(T)$  a  $\mathbf{y}_{(T+\tau)}(T) - \hat{\mathbf{y}}_{(T+\tau)}(T)$  jsou nekorelované náhodné vektory, další informace nalezneme v Lütkepohl (2005, odstavec 10.5).

**Poznámka 1.8.1.** Dhrymes (1994, odstavec 2.2) uvádí asymptotické vlastnosti předpovědí v nedynamických i dynamických modelech a provádí jejich srovnání.  $\diamond$

## 1.9 Výpočetní software

Při praktické aplikaci prezentovaných teoretických konceptů, především při vlastním odhadu parametrů, testování hypotéz či konstrukci předpovědí, bývá obvykle používán ekonometrický software, který do značné míry usnadňuje samotné numerické výpočty.

Jedním z příkladů dostupných programových řešení je uživatelsky přátelský EViews 6.0 vyvíjený americkou firmou QMS (Quantitative Micro Software), Irvine (CA), který použijeme při následné praktické analýze ekonometrického modelu. Obsahuje všechny v praxi běžně používané odhadové procedury (typu OLS, GMM, 2SLS, 3SLS, FIML, včetně jejich případných iteračních modifikací), umožňuje konstruovat předpovědi, a to jak deterministickým, tak stochastickým přístupem (viz odstavec 3.3). Testy exogenity, specifikace či test uvedený v poznámce 1.4.2, zde však nejsou implementovány v přímé podobě, je tak nutné je počítat ad hoc, což ovšem vzhledem k jejich charakteru není příliš obtížné. Poznamenejme, že nejnovější verze EViews 7.0 navíc disponuje odhady  $k$ -té třídy, LIML-odhady a je rovněž explicitně schopna provádět testování exogenity<sup>1</sup>.

Výpočetní prostředí pro soustavy simultánních ekonometrických rovnic je samozřejmě obsaženo v mnoha dalších softwarových produktech, např. SAS, Stata, apod.

---

<sup>1</sup>Informace převzaty ze stránek výrobce [www.eviews.com/EViews7/ev7whatsnew.html](http://www.eviews.com/EViews7/ev7whatsnew.html), navštíveno dne 16. března 2010.

## Kapitola 2

# Základní pojmy životního pojištění

V této kapitole se stručně seznámíme s vybranými pojmy z oblasti životního pojištění, podrobnějším zdrojem informací může být např. Cipra (2006a) nebo Cipra (2006b), podle nichž je tato kapitola zpracována.

V kontextu soukromého životního pojištění (*life insurance*) se obvykle setkáváme s dlouhodobějšími pojistnými smlouvami, jejichž hlavním úkolem je zajistit finanční podporu v případě ztráty příjmu, uplatňují se tu však i jiné motivační prvky (např. investiční).

Účastníci životního pojištění se označují jako:

- **pojistitel**: je právnická osoba (zpravidla pojišťovna), která je v souladu se zákonem oprávněna provozovat pojišťovací činnost;
- **pojistník**: je fyzická nebo právnická osoba, která s pojistitelem uzavřela pojistnou smlouvu, mezi její povinnosti patří placení pojistného;
- **pojištěný (pojištěnec, účastník)**: je fyzická osoba, na jejíž život a zdraví se pojištění vztahuje;
- **oprávněná osoba**: je fyzická nebo právnická osoba, které v důsledku pojistné události vznikne právo na pojistné plnění; pojistníkem, pojištěným a oprávněnou osobou může být jeden jedinec;
- **obmyšlený**: je fyzická či právnická osoba uvedená pojistníkem v pojistné smlouvě, které vznikne právo na pojistné plnění v případě smrti pojištěného (pokud není v pojistné smlouvě obmyšlený explicitně uveden, postupuje se dle zákonných norem).

**Pojistné** můžeme rozdělit do několika skupin:

- **jednorázové pojistné**: je stanoveno na celou dobu sjednaného pojištění, tj. platí se najednou při uzavření pojistné smlouvy;
- **běžné pojistné**: je stanoveno za *pojistné období*, tj. platí se opakovaně v pravidelných splátkách obvykle na počátku jednotlivých pojistných období;

- *nettopojistné (ryzí pojistné)*: je vypočteno tak, aby pojistiteli v průměru pokrylo vyplácená pojistná plnění;
- *bruttopojistné (hrubé pojistné)*: je nettopojistné rozšířené o částky vydané na pokrytí *správních nákladů* pojistitele a případných nepříznivých škodních výchylek formou *bezpečnostní přírážky*;
- *valorizované pojistné*: je pojistné navyšované se souhlasem pojistníka vzhledem k vývoji inflace.

**Poznámka 2.0.1.** V případě neplacení běžného pojistného pojistníkem trvá v klasickém případě pojištění i nadále do konce sjednané doby s *redukovanou pojistnou částkou* či *redukováným důchodem*. Ve flexibilním životním pojištění (umožňuje pružnou realizaci nejrůznějších změn) lze za stanovených podmínek vynechat několik splátek pojistného (tzv. pojistné prázdniny). V některých specifických situacích (např. invalidita pojistníka) je rovněž možné pojistníka zprostit povinnosti platit pojistné. ◇

**Pojistná událost** je nahodilá skutečnost, která je blíže specifikována v pojistné smlouvě a s níž je spojen vznik povinnosti pojistitele poskytnout pojistné plnění.

**Pojistné plnění** se vyplácí ve formě: *jednorázové pojistné částky, důchodu* (tj. jednorázová pojistná částka se po dohodě s klientem rozloží do anuity; rozlišuje se mezi *životním důchodem*, jehož výplata je podmíněna tím, že je pojištěný naživu, a *jistým důchodem*, jenž bývá vyplácen po stanovenou dobu bez ohledu na předchozí podmínku) a *zproštění od placení pojistného* (institut umožňující pojistiteli ve specifických situacích převzít povinnost pojistníka platit pojistné, např. invalidita pojistníka, viz výše).

**Pojistná doba** je doba, na kterou bylo pojištění sjednáno, obvykle se setkáváme s *dočasným pojištěním* (pojistná doba je předem smluvně omezena), s *trvalým pojištěním* (pojistná doba není smluvně předem omezena) a s *pojištěním s odkladem* (je sjednána doba odkladu, o kterou je odložena povinnost pojistitele platit pojistné plnění).

V rámci životního pojištění osob lze sjednat:

- **pojištění pro případ smrti**: pojistnou událostí je smrt pojištěného;
- **pojištění pro případ dožití**: pojistnou událostí je dožití sjednaného věku pojištěného;
- **smíšené pojištění**: pojistnou událostí je smrt pojištěného či dožití daného věku pojištěným, podle toho, co nastane dříve;
- **důchodové pojištění**: jedná se o speciální případ pojištění pro případ dožití s pravidelně se opakujícím pojistným plněním ve formě výplaty důchodu.

Odlišné členění životního pojištění je založeno na tom, zda pojistný produkt obsahuje či neobsahuje spořivou (investiční) složku vytvářenou z pojistného a nazývanou *rezerva pojistného*:

- **rezervotvorné životní pojištění** (kapitálové životní pojištění);
- **nerezervotvorné životní pojištění**: většinou se jedná o produkty dočasného pojištění pro případ smrti.

Uvedme některé další důležité pojmy vztahující se k životnímu pojištění.

- **Pojistný kmen** je soubor pojistných smluv obdobného typu.
- **Technická úroková míra (TÚM)** je úroková míra používaná k oceňování systémů finančních toků v životním pojištění (tj. pro kalkulaci pojistného, technických rezerv, apod.). Představuje takové zhodnocení pojistného, na které má klient smluvní nárok (skutečné zhodnocení však může být vyšší). Bývá volena konzervativně, v ČR je její výše regulována zákonnou normou. Nízká (vysoká) TÚM zvyšuje (snižuje) pojistné sazby.
- **Odkup (odbytné)** je vrácení části rezervy pojistného životního pojištění u rezervotvorných pojištění pojistníkovi, který o toto požádá. V praxi se rozlišuje mezi zánikem pojištění bez výplaty náhrady klientovi a mezi zrušením pojištění (na žádost klienta), které je spojeno s výplatou odkupného.
- **Podíl na zisku pojišťovny** je rozdělování toho zisku životní pojišťovny jejím klientům, který pojišťovna dosahuje díky rozdílu mezi kalkulovaným stavem (tzv. *výpočetní podklady prvního řádu, statutární výpočetní podklady*) a skutečným stavem (tzv. *výpočetní podklady druhého řádu*) příslušných parametrů, a to zejména míry zisku, úmrtnosti, stornovosti a správním nákladům.
- **Opce v životním pojištění** znamená možnost volby klienta z několika variant v různých fázích platnosti pojistné smlouvy.
- **Sazebník životní pojišťovny** uvádí pro jednotlivé pojistné produkty výši bruttopojistného, přičemž se v úvahu bere pohlaví pojištěného, vstupní věk pojištěného a pojistná doba.

**Správní náklady** pojistitele můžeme rozčlenit do následujících skupin:

- **náklady spojené se sjednáním pojištění  $\alpha$** : započítávají se jako procenta z pojistné částky nebo z ročního důchodu; lze rozlišovat mezi *jednorázovými počátečními (získávacími, akvizičními) náklady  $\alpha^Z$* , které se vynakládají ihned při uzavření pojistné smlouvy (získávací provize pojistným zprostředkovatelům, vystavení pojistné smlouvy, lékařská prohlídka, apod.), a *provizními náklady  $\alpha^P$* , ze kterých každoročně plynou udržovací provize pro pojistné zprostředkovatele;
- **běžné správní náklady  $\beta$** : jsou každoroční náklady během trvání pojištění (pojistiteli zabezpečují personální náklady, administrativu, provoz výpočetní techniky, apod.), započítávají se jako procenta z pojistné částky nebo z ročního důchodu; je-li doba placení pojistného kratší než pojistná doba rozdělují se na běžné správní náklady  $\beta_1$  *během placení pojistného* a na běžné správní náklady  $\beta_2$  *během neplacení pojistného*;



- *inkasní náklady*  $\gamma$ : jsou náklady spojené s inkasem (běžného) pojistného, započítávají se jako procenta z ročního bruttopojistného;
- *náklady při výplatě důchodu*  $\delta$ : započítávají se jako procenta z ročního důchodu;
- *jednotná správní přírážka*  $\varepsilon$ : u některých pojišťoven v sobě soustředí všechny typy předchozích správních nákladů, započítává se jako procenta z bruttopojistného.

**Technické rezervy** jsou rezervy finančních prostředků, které musí podle zákona pojišťovna vytvářet (zákonné normy obvykle stanovují typ i přiměřenou výši technických rezerv podléhající schválení příslušným dozorovým orgánem), k plnění závazků z pojišťovací činnosti, které jsou pravděpodobné nebo jisté, ale není jistá jejich výše nebo okamžik jejich vzniku. Technické rezervy představují podstatnou součást pasiv každé pojišťovny a o každé se účtuje odděleně od jiných závazků pojistitele. Aktiva, jejichž zdrojem jsou technické rezervy, podléhají přísným restrikcím tak, aby skladba jejich finanční alokace odpovídala zásadám bezpečnosti, diverzifikace, rentability a likvidity. V rámci legislativy ČR je specifikováno několik skupin technických rezerv, tyto jsou uvedeny např. v Cípra (2006b, kapitola 11).

**Zajištěním (reinsurance)** rozumíme převod určité části rizika, které na sebe převzal pojistitel od pojištěných, na jiný subjekt pojišťovacího trhu, který již obvykle není k pojištěným vázán žádnými smluvními závazky. Pojišťovna (*prvopojistitel, cedent*) v tomto případě postoupí (*ceduje*) část rizika převzatého od pojištěných na zajišťovnu (*zajistitele, cesionáře*). Část rizika, které si ponechá prvopojistitel, označujeme jako *vlastní vrub (maximum)*. Za příslušný převod rizika musí prvopojistitel postoupit část inkasovaného pojistného zajistiteli. Vztah prvopojistitele a zajišťovny je tedy v určitém smyslu analogický vztahu pojištěného a pojistitele. Problematika zajištění je poměrně obsáhlá, podrobnosti lze nalézt např. v Cípra (2006b, kapitola 19).

# Kapitola 3

## Ekonometrický model životní pojišťovny

### 3.1 Formulace modelu

V této části práce ukážeme, jakým způsobem lze využít dynamických ekonometrických soustav (lineárních) simultánních rovnic k analýze finančních toků v životní pojišťovně.

V období let 1996 až 2005 sledujeme životní pojišťovnu operující na území České republiky (data byla získána kompilací z Cipra (1998) a Cipra (2008, příklad 7.4.3)). Konkrétně máme k dispozici tyto proměnné ( $t = 1996, \dots, 2005$ ):

$N_t$  - počet nově uzavřených pojistných smluv v  $t$ -tém roce (v ks);

$I_t$  - technická úroková míra (TÚM) v roce  $t$  (v % p.a.);

$IEF_t$  - eficientní úroková míra v roce  $t$  (v % p.a.), tj. míra zisku z investiční činnosti pojišťovny;

$G_t$  - koeficient růstu průměrné pojistné částky na jednu pojistnou smlouvu v roce  $t$  ve srovnání s předchozím obdobím (v % p.a.);

$EX_t$  - počet ukončených pojistných smluv (pojistná událost, odkup, atp.) v  $t$ -tém roce (v ks);

$KMEN_t$  - pojistný kmen v  $t$ -tém roce (v ks);

$P_t$  - předepsané pojistné za  $t$ -tý rok (v tis. Kč);

$CS_t$  - pojistné plnění (včetně odkupů) v roce  $t$  (v tis. Kč);

$V_t$  - technická rezerva pojistného v  $t$ -tém roce (v tis. Kč);

$EAC_t$  - počáteční správní náklady za  $t$ -tý rok (v tis. Kč);

$EAD_t$  - běžné správní náklady za  $t$ -tý rok (v tis. Kč);

$RE_t$  - výsledek zajištění v  $t$ -tém roce (v tis. Kč), tj. zajistné plnění plus provize od zajistitele po odečtení zajistného;

$PROF_t$  - investiční výnos pojišťovny za  $t$ -tý rok (v tis. Kč);

$PROFS_t$  - podíl na zisku pojišťovny v  $t$ -tém roce (v tis. Kč);

$R_t$  - pojistný výsledek životní pojišťovny (celkové výnosy minus celkové náklady) za  $t$ -tý rok (v tis. Kč).

Vzhledem k avizovanému dynamickému modelování je možné zavést několik dalších proměnných, které model přiblíží skutečnosti (např. pomocí proměnné  $G_t$  lze určitým způsobem zohlednit míru inflace). Jedná se o následující proměnné:

$NG_t$  - proměnná  $N_t$  upravená o růst průměrné pojistné částky,

$$NG_t = (1 + G_1) \dots (1 + G_{t-1}) \sqrt{1 + G_t} N_t = \kappa_t N_t,$$

kde  $\kappa_t = \sqrt{1 + G_t} \prod_{i=1}^{t-1} (1 + G_i)$ ;

$KMENG_t$  - pojistný kmen  $KMEN_t$  upravený o růst průměrné pojistné částky,

$$KMENG_t = \kappa_t \cdot \frac{KMEN_t + KMEN_{t-1}}{2};$$

$VI_t$  - technické rezervy  $V_t$  upravené pomocí technické úrokové míry  $I_t$ ,

$$VI_t = (1 + I_t) \cdot \frac{V_t + V_{t-1}}{2};$$

$VIEF_t$  - technické rezervy  $V_t$  upravené pomocí eficientní úrokové míry  $IEF_t$ ,

$$VIEF_t = (1 + IEF_t) \cdot \frac{V_t + V_{t-1}}{2}.$$

Uvedme definiční vztahy (identity) pro proměnné  $KMEN_t$  a  $R_t$

$$KMEN_t = KMEN_{t-1} + N_t - EX_t, \quad (3.1.1)$$

$$\begin{aligned} R_t &= (P_t + RE_t + PROF_t) - \\ &- (CS_t + (V_t - V_{t-1}) + EAC_t + EAD_t + PROFS_t). \end{aligned} \quad (3.1.2)$$

Interpretace druhé prezentované identity (3.1.2) je v souladu s výše uvedenou specifikací proměnné  $R_t$ , neboť členy v první závorce na pravé straně rovnosti představují celkové výnosy pojišťovny v roce  $t$  a členy v druhé závorce tvoří celkové náklady pojišťovny v roce  $t$ .

Tabulka 3.1.1: Data k ekonometrickému modelu finančních toků v životní pojišťovně.

rok	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
$I_t^{(1)}$	4.00	4.00	4.00	4.00	4.00	4.00	3.00	3.00	2.40	2.40
$IEF_t^{(1)}$	10.00	10.00	9.00	8.00	8.00	7.00	7.50	7.00	6.00	6.50
$G_t^{(1)}$	2.00	2.00	2.00	2.50	2.50	3.50	2.50	3.00	3.50	3.00
$N_t^{(2)}$	14385	10098	12105	13951	16029	17787	15209	15855	13980	13567
$NG_t^{(2)}$	14528	10402	12719	14989	17652	20175	17769	19033	17327	17362
$KMEN_t^{(2)}$	116801	123095	131187	140672	151925	164443	174344	186189	196172	206883
$KMENG_t^{(2)}$	111964	123564	133594	146041	161111	179424	197901	216394	236954	257895
$EX_t^{(2)}$	2505	3804	4013	4466	4776	5269	5308	4010	3997	2856
$P_t^{(3)}$	180985	278256	303382	345692	380456	426910	450212	480945	501098	576054
$CS_t^{(3)}$	56091	75899	84214	97774	109029	124597	130055	151007	143997	139807
$V_t^{(3)}$	859086	1006740	1164741	1347763	1546061	1769684	1870935	1899566	1959002	2027891
$VI_t^{(3)}$	843719	970230	1129170	1306502	1504788	1724187	1874919	1941808	1975587	2041289
$VIEF_t^{(3)}$	892395	1026204	1183457	1356752	1562665	1773924	1956833	2017218	2045041	2123021
$EAC_t^{(3)}$	15705	14699	15623	18517	19652	22018	25955	28551	29123	34942
$EAD_t^{(3)}$	5052	8624	8513	9723	9807	10882	12922	14707	14521	14235
$RE_t^{(3)}$	2598	1860	1926	2234	2366	2680	3012	2944	2895	3192
$PROF_t^{(3)}$	40256	68291	73512	80455	86626	102813	110557	119444	130225	135782
$PROFS_t^{(3)}$	34599	60047	60485	70387	71632	87391	92985	98006	120999	123098
$R_t^{(3)}$	16757	41484	51984	48958	61030	63892	200613	282431	266142	334057

<sup>(1)</sup> - % p.a., <sup>(2)</sup> - počet pojistných smluv v ks, <sup>(3)</sup> - tis. Kč.  $KMEN_{1995} = 104921$  ks a  $V_{1995} = 763451$  tis. Kč.

Nyní přistoupíme k formulaci vlastních rovnic ekonometrického modelu životní pojišťovny. Pro  $t = 1996, \dots, 2005$  pracujeme se vztahy

$$P_t = a_1 + a_2 KMENG_t + \varepsilon_t^P, \quad (3.1.3)$$

$$CS_t = b_1 + b_2 KMENG_t + b_3 V_{t-1} + b_4 EX_t + \varepsilon_t^{CS}, \quad (3.1.4)$$

$$V_t = c_1 + c_2 V_{t-1} + c_3 P_t + c_4 CS_t + \varepsilon_t^V, \quad (3.1.5)$$

$$EAC_t = d_1 + d_2 KMENG_t + d_3 NG_t + d_4 P_t + \varepsilon_t^{EAC}, \quad (3.1.6)$$

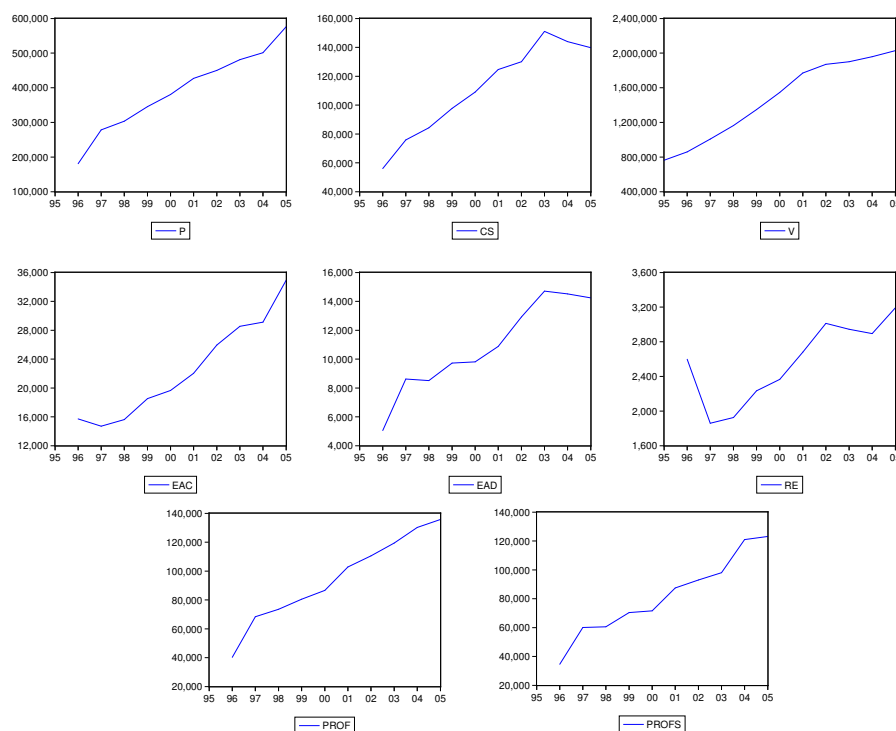
$$EAD_t = e_1 + e_2 KMENG_t + e_3 NG_t + e_4 CS_t + \varepsilon_t^{EAD}, \quad (3.1.7)$$

$$RE_t = f_1 + f_2 KMENG_t + f_3 NG_t + f_4 EX_t + \varepsilon_t^{RE}, \quad (3.1.8)$$

$$PROF_t = g_1 + g_2 (VIEF_t - VI_t) + g_3 P_t + g_4 CS_t + \varepsilon_t^{PROF}, \quad (3.1.9)$$

$$PROFS_t = h_1 + h_2 PROF_t + \varepsilon_t^{PROFS}. \quad (3.1.10)$$

Je tedy zřejmé, že právě prezentovaný model finančních toků v životní pojišťovně skládající se ze dvou identit (3.1.1) - (3.1.2) a z osmi rovnic (3.1.3) - (3.1.10) je dynamickou ekonometrickou soustavou lineárních simultánních rovnic, ve které intercept a proměnné  $EX_t$ ,  $NG_t$  a  $KMENG_t$  vystupují jako striktně exogenní proměnné a proměnná  $V_{t-1}$  je exogenní ve smyslu predeterminovanosti. Není obtížné ověřit, že všechny rovnice modelu (3.1.3) - (3.1.10) splňují nutnou podmínku identifikace (1.2.7).



**Obrázek 3.1.1:** Průběh vysvětlovaných proměnných v modelu životní pojišťovny.

## 3.2 Odhad parametrů modelu

K odhadu parametrů uvedeného modelu (3.1.3) - (3.1.10) použijeme vybrané odhadové procedury, se kterými jsme se seznámili v odstavci 1.4. Konkrétně jsme se rozhodli aplikovat 2SLS-odhad a 3SLS-odhad. Shrnutí jejich vlastností nalezneme v odstavcích 1.4.4 a 1.4.5, odtud je také zřejmé, že apriori preferujeme druhý ze jmenovaných přístupů.

Poznamenejme, že výpočetní implementace úlohy byla provedena v programu EViews 6.0.

V tabulce 3.2.1 jsou uvedeny výsledné odhady jednotlivých parametrů modelu získané pomocí 2SLS a 3SLS přístupu (včetně směrodatných odchylek a t-poměrů, tj. podíl odhadu parametru a jeho směrodatné odchylky).

	<b>2SLS</b>	sm. odch.	t-poměr	<b>3SLS</b>	sm. odch.	t-poměr
$a_1$	-15004.49	36183.52	-0.414677	-13366.26	32166.17	-0.415538
$a_2$	2.308440	0.198053	11.65565	2.299158	0.175986	13.06440
$b_1$	2804.041	22797.46	0.122998	4415.197	12332.71	0.358007
$b_2$	-0.040380	0.324772	-0.124334	-0.071585	0.160571	-0.445813
$b_3$	0.071970	0.036750	1.958362	0.074938	0.017748	4.222425
$b_4$	3.283935	4.266427	0.769715	3.207209	2.124543	1.509599
$c_1$	48240.63	156955.9	0.307351	73028.33	108436.0	0.673470
$c_2$	-0.087499	0.719323	-0.121640	0.084245	0.437638	0.192499
$c_3$	0.979505	1.241547	0.788940	0.895680	0.772533	1.159406
$c_4$	11.11656	8.626938	1.288587	8.999220	5.619008	1.601567
$d_1$	-3202.897	2649.538	-1.208851	-2417.507	1677.186	-1.441407
$d_2$	0.221252	0.048124	4.597560	0.226518	0.017604	12.86709
$d_3$	0.116809	0.213387	0.547403	0.049720	0.127967	0.388535
$d_4$	-0.038884	0.020965	-1.854717	-0.040485	0.006863	-5.899399
$e_1$	2936.058	476.7411	6.158601	2899.449	313.1464	9.259084
$e_2$	0.005446	0.005135	1.060616	0.003401	0.002502	1.358926
$e_3$	-0.358958	0.045248	-7.933205	-0.357437	0.025606	-13.95897
$e_4$	0.115193	0.009711	11.86186	0.118546	0.004240	27.96079
$f_1$	989.1966	427.2213	2.315420	842.7909	277.9763	3.031880
$f_2$	0.003985	0.002011	1.981831	0.003953	0.001288	3.068665
$f_3$	0.095550	0.036377	2.626649	0.098384	0.016635	5.914337
$f_4$	-0.163215	0.088218	-1.850120	-0.137318	0.042662	-3.218767
$g_1$	-12155.31	8817.446	-1.378552	-10625.05	5478.173	-1.939525
$g_2$	0.221836	0.251013	0.883763	0.117332	0.139907	0.838644
$g_3$	0.236061	0.059419	3.972858	0.213651	0.033254	6.424753
$g_4$	0.004041	0.227482	0.017763	0.128074	0.146728	0.872867
$h_1$	-5088.929	4595.530	-1.107365	-6080.967	4084.174	-1.488910
$h_2$	0.918306	0.046409	19.78729	0.928771	0.041221	22.53169

**Tabulka 3.2.1:** 2SLS-odhady, resp. 3SLS-odhady parametrů rovnic (3.1.3) - (3.1.10) v ekonometrickém modelu finančních toků v životní pojišťovně (včetně příslušných směrodatných odchylek a t-poměrů).

Na základě výsledků odhadu lze provést Hausmanův test specifikace modelu, o kterém pojednává odstavce 1.6. Hodnota Waldovy testové statistiky (1.6.1) je  $W = 1.73276$  s  $p$ -hodnotou 0.999999999999313, nulovou hypotézou tudíž nemůžeme na standardní 5% hladině spolehlivosti zamítnout, tj. neprokázali jsme nevhodnost specifikace modelu ve smyslu uvedeném v citovaném odstavci.

Tabulka 3.2.2 obsahuje hodnoty koeficientů determinace určené na základě 2SLS-odhadu, resp. 3SLS-odhadu parametrů rovnic (3.1.3) - (3.1.10) ekonometrického modelu finančních toků v životní pojišťovně. Vidíme, že koeficienty jsou v obou případech dostatečně vysoké. Dále uvádíme hodnoty testových statistik  $Sa$  (viz poznámka 1.4.2, vztah (1.4.30)) s odpovídajícími  $p$ -hodnotami. Je zřejmé, že v žádné rovnici na 5% hladině spolehlivosti nezamítáme nulovou hypotézu tohoto testu, tj. že exogenní proměnné a residuální složka splňují v dané rovnici podmínku ortogonalit (současné nekorelovanosti).

rovnice	${}_{2SLS}R^2$	${}_{3SLS}R^2$	$Sa (df)$	$p$ -hodnota
(3.1.3)	0.94439	0.94437	6.61938 (3)	0.08507
(3.1.4)	0.97457	0.97445	0.18781 (1)	0.66474
(3.1.5)	0.97202	0.97859	2.69061 (1)	0.10094
(3.1.6)	0.97274	0.97114	0.15889 (1)	0.69018
(3.1.7)	0.99631	0.99581	2.81908 (1)	0.09315
(3.1.8)	0.86044	0.85676	2.11896 (1)	0.14549
(3.1.9)	0.98167	0.98652	0.78603 (1)	0.37530
(3.1.10)	0.97998	0.97976	6.68263 (3)	0.08273

**Tabulka 3.2.2:** Ve druhém a třetím sloupci tabulky jsou uvedeny koeficienty determinace určené na základě 2SLS-odhadu, resp. 3SLS-odhadu parametrů rovnic modelu (3.1.3) - (3.1.10). Ve čtvrtém sloupci je prezentována hodnota testové statistiky  $Sa$  (viz (1.4.30)) pro danou rovnici a příslušný počet stupňů volnosti  $df$ ,  $p$ -hodnota tohoto testu je obsažena v posledním sloupci tabulky.

Právě provedený odhad ekonometrického modelu životní pojišťovny (3.1.3) - (3.1.10) nám poskytuje mnoho relevantních informací, např.

- z 2SLS-odhadu, resp. 3SLS-odhadu rovnice (3.1.3), tj.

$$P_t = -15004.48919 + 2.30844 \cdot KMENG_t,$$

resp.  $P_t = -13366.25500 + 2.29916 \cdot KMENG_t,$

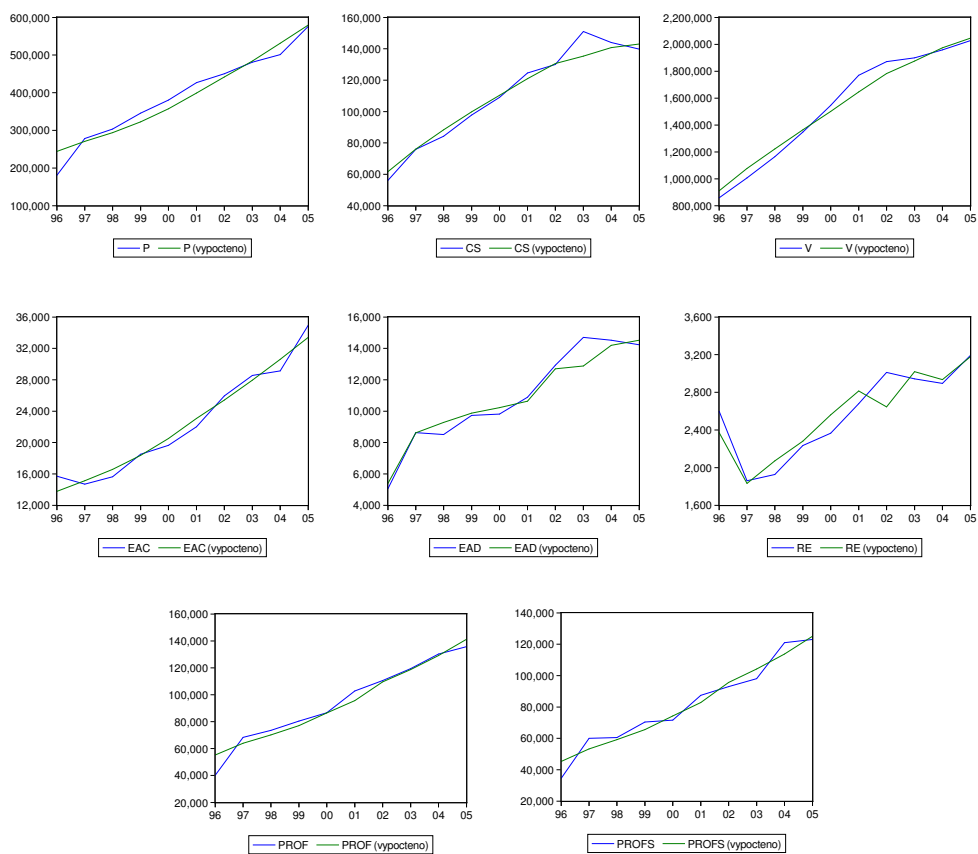
vyplývá, že průměrné roční pojistné je 2308.44 Kč, resp. 2299.16 Kč na jednu pojistnou smlouvu a zvyšuje se ročně v souladu s růstem průměrné pojistné částky;

- podle 2SLS-odhadu, resp. 3SLS-odhadu rovnice (3.1.10), tj.

$$PROFS_t = -5088.92950 + 0.91831 \cdot PROF_t,$$

resp.  $PROFS_t = -6080.96665 + 0.92877 \cdot PROF_t,$

pojišťovna v průměru přiděluje klientům 91.83 %, resp. 92.88 % svého investičního výnosu jako podíl na zisku.



**Obrázek 3.2.1:** Průběh vysvětlovaných proměnných a jejich (simultánně) vypočtených hodnot na základě 3SLS-odhadu ekonometrického modelu životní pojišťovny.



### 3.3 Použití bootstrapu

*Metoda bootstrap* byla poprvé představena na konci sedmdesátých let 20. století v článku Efron (1979), jedná se tedy o poměrně novou statistickou metodu, která k analýze dat intenzivním způsobem využívá výpočetní techniku. Základní přehled pojmů a postupů bootstrapu nabízí např. Prášková (2004), zde je rovněž možné najít odkazy na další relevantní zdroje.

Také v rámci teorie ekonometrických soustav simultánních rovnic našel bootstrap své uplatnění, obsáhlejší práci na toto téma je Freedman (1984), praktickou implementaci a výpočetní algoritmy uvádí Faire (2004, kapitola 9). Nyní se ve stručnosti seznámíme se základními výsledky a postupy aplikace bootstrapu v této konkrétní situaci.

Uvažujme strukturální tvar dynamické soustavy simultánních rovnic (1.7.2)

$$\mathbf{y}_t \mathbf{\Gamma} + \sum_{k=1}^p \mathbf{y}_{(t-k)} \cdot \mathbf{\Phi}_k + \mathbf{x}_t \mathbf{B} + \boldsymbol{\varepsilon}_t = \mathbf{0}, \quad t = 1, \dots, T, \quad (3.3.1)$$

přičemž k již zavedeným předpokladům **(P1)** - **(P3)**, resp. **(A1)** - **(A3)** a invertibilitě matice  $\mathbf{\Gamma}$  navíc požadujeme například (ve shodě s Freedman (1984)), aby  $\mathbf{x}_t$  ( $x_{t1} = 1$ ) a  $\boldsymbol{\varepsilon}_t$  byly nezávislé stejně rozdělené náhodné vektory, jejichž složky mají konečné čtvrté momenty, aby uvedená soustava byla stabilní (viz odstavec 1.7) a aby  $\{\mathbf{y}_t\}$  byla slabě stacionární posloupností, tj.

$$\mathbf{E}(\mathbf{y}'_t) = \boldsymbol{\mu},$$

$$\text{cov}(\mathbf{y}'_s, \mathbf{y}'_t) = \mathbf{E}(\mathbf{y}'_s - \boldsymbol{\mu})(\mathbf{y}'_t - \boldsymbol{\mu})' = \text{cov}(\mathbf{y}'_{(s+h)}, \mathbf{y}'_{(t+h)}), \quad \forall h.$$

Další podrobnosti uvádí Freedman (1984).

V kontextu problematiky ekonometrických soustav simultánních rovnic (ve tvaru (3.3.1)) se nejčastěji setkáváme s tzv. *residuálním bootstrapem*, který se při známých hodnotách striktně exogenních proměnných  $\mathbf{x}_1, \dots, \mathbf{x}_T$  a při daných  $\mathbf{y}_{(-p+1)}, \dots, \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_T$  řídí následujícím algoritmem. Ještě poznamenejme, že středem našeho zájmu je funkce neznámých parametrů  $q = q(\mathbf{\Gamma}, \mathbf{B}, \mathbf{\Phi}_1, \dots, \mathbf{\Phi}_p)$ , resp. její odhad  $\hat{q}$ :

1. Na základě  $\hat{\mathbf{\Gamma}}$ ,  $\hat{\mathbf{B}}$  a  $\hat{\mathbf{\Phi}}_k$ ,  $k = 1, \dots, p$ , konzistentních odhadů parametrů  $\mathbf{\Gamma}$ ,  $\mathbf{B}$  a  $\mathbf{\Phi}_k$  modelu (3.3.1) (zpravidla aplikujeme odhadové procedury typu 2SLS či 3SLS) vypočteme *odhadnutá residua*

$$\hat{\boldsymbol{\varepsilon}}_t = -\mathbf{y}_t \hat{\mathbf{\Gamma}} - \sum_{k=1}^p \mathbf{y}_{(t-k)} \cdot \hat{\mathbf{\Phi}}_k - \mathbf{x}_t \hat{\mathbf{B}}, \quad t = 1, \dots, T,$$

tato vycentrujeme, obdržíme tak posloupnost  $\hat{\boldsymbol{\varepsilon}}_1 - \bar{\boldsymbol{\varepsilon}}, \dots, \hat{\boldsymbol{\varepsilon}}_T - \bar{\boldsymbol{\varepsilon}}$ , ve které označujeme  $\bar{\boldsymbol{\varepsilon}} = \frac{1}{T} \sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}_t$ .

2. Z výše uvedené centrované posloupnosti získáme náhodným výběrem s vracením tzv. *bootstrapová residua*  $\boldsymbol{\varepsilon}_1^*, \dots, \boldsymbol{\varepsilon}_T^*$ , na jejichž základě spočteme

$$\mathbf{y}_t^* = -\sum_{k=1}^p \mathbf{y}_{(t-k)}^* \cdot \hat{\mathbf{\Phi}}_k \hat{\mathbf{\Gamma}}^{-1} - \mathbf{x}_t \hat{\mathbf{B}} \hat{\mathbf{\Gamma}}^{-1} - \boldsymbol{\varepsilon}_t^* \hat{\mathbf{\Gamma}}^{-1}, \quad t = 1, \dots, T,$$

přičemž pokládáme  $\mathbf{y}_{(-p+1)}^* = \mathbf{y}_{(-p+1)}, \dots, \mathbf{y}_0^* = \mathbf{y}_0$ .

3. Konzistentně odhadneme model (3.3.1), ve kterém namísto  $\mathbf{y}_{(-p+1)\cdot}, \dots, \mathbf{y}_{0\cdot}, \mathbf{y}_{1\cdot}, \dots, \mathbf{y}_T$  uvažujeme sadu hodnot  $\mathbf{y}_{(-p+1)\cdot}^*, \dots, \mathbf{y}_{0\cdot}^*, \mathbf{y}_{1\cdot}^*, \dots, \mathbf{y}_T^*$ . Na základě tohoto odhadu vypočítáme  $\hat{q}^* = \hat{q}^* \left( \hat{\Gamma}^*, \hat{\mathbf{B}}^*, \hat{\Phi}_1^*, \dots, \hat{\Phi}_p^* \right)$ , tj. hodnotu *bootstrapové verze odhadu*  $\hat{q}$ .
4. Druhý až třetí krok algoritmu opakujeme  $B$ -krát,  $B \in \mathbb{N}$  je voleno dostatečně velké.

Po ukončení algoritmu bychom měli mít k dispozici hodnoty  $\hat{q}_1^*, \dots, \hat{q}_B^*$ , na jejichž základě definujeme  $\bar{q}^*$  *bootstrapový odhad*  $q$  jako

$$\bar{q}^* = \frac{1}{B} \sum_{i=1}^B \hat{q}_i^* \quad (3.3.2)$$

s *bootstrapovým odhadem rozptylu*

$$\widehat{\text{var}}(\bar{q}^*) = \frac{1}{B-1} \sum_{j=1}^B \left( \hat{q}_j^* - \frac{1}{B} \sum_{i=1}^B \hat{q}_i^* \right)^2. \quad (3.3.3)$$

Velmi často se metoda bootstrap využívá ke konstrukci intervalů spolehlivosti. V této souvislosti představíme studentizované intervaly spolehlivosti, které v uvažovaném kontextu patří mezi nejvíce frekventované.

Mějme studentizovanou statistiku

$$R = \frac{q - \hat{q}}{\sqrt{\widehat{\text{var}}(\hat{q})}}$$

a její bootstrapový protějšek

$$R^* = \frac{\hat{q}^* - \hat{q}}{\sqrt{\widehat{\text{var}}(\hat{q}^*)}},$$

odhad  $\widehat{\text{var}}(\hat{q}^*)$  se hledá standardními statistickými postupy. Alternativou je rovněž opakované použití bootstrapu. Pro každý odhad  $\hat{q}^*$ , který jsme obdrželi pomocí výše uvedeného algoritmu na základě  $\mathbf{y}_{(-p+1)\cdot}^*, \dots, \mathbf{y}_{0\cdot}^*, \mathbf{y}_{1\cdot}^*, \dots, \mathbf{y}_T^*$ , použijeme celý zmiňovaný algoritmus s  $B_1$  opakováními znovu, nicméně tentokrát v modelu (3.3.1) namísto  $\mathbf{y}_{(-p+1)\cdot}, \dots, \mathbf{y}_{0\cdot}, \mathbf{y}_{1\cdot}, \dots, \mathbf{y}_T$  uvažujeme sadu hodnot  $\mathbf{y}_{(-p+1)\cdot}^*, \dots, \mathbf{y}_{0\cdot}^*, \mathbf{y}_{1\cdot}^*, \dots, \mathbf{y}_T^*$ . Tudíž v rámci každého z  $B$  opakování algoritmu pro výpočet  $\hat{q}^*$  spouštíme totožný algoritmus pro modifikovaný model s  $B_1$  opakováními; celkem tak generujeme  $B \cdot B_1$  bootstrapových residuí.

Je-li  $H$  distribuční funkce  $R$  a  $\gamma_p$  je příslušný  $p$ -kvantil, potom je interval spolehlivosti pro  $q$  s koeficientem spolehlivosti  $(1 - \alpha)$  definován jako

$$CI = \left( \hat{q} - \gamma_{(1-\alpha/2)} \sqrt{\widehat{\text{var}}(\hat{q})}, \hat{q} - \gamma_{\alpha/2} \sqrt{\widehat{\text{var}}(\hat{q})} \right). \quad (3.3.4)$$

*Bootstrapový interval spolehlivosti* s koeficientem spolehlivosti  $(1 - \alpha)$  je tvaru

$$CI^* = \left( \hat{q} - \gamma_{(1-\alpha/2)}^* \sqrt{\widehat{\text{var}}(\hat{q})}, \hat{q} - \gamma_{\alpha/2}^* \sqrt{\widehat{\text{var}}(\hat{q})} \right), \quad (3.3.5)$$

kde  $\gamma_p^*$  je  $p$ -kvantil distribuční funkce  $H^*$  statistiky  $R^*$  spočtený jako  $\gamma_p^* = R_{(\lfloor Bp \rfloor)}^*$ ,  $[\cdot]$  označuje dolní celou část, tj. jedná se o výběrový kvantil spočtený v uspořádaném výběru  $R_{(1)}^*, \dots, R_{(B)}^*$ .

Jak konstatuje Lütkepohl (2005, strana 713),

$$\sqrt{T}(\hat{q} - q) \quad \text{a} \quad \sqrt{T}(\hat{q}^* - \hat{q})$$

mají za platnosti obecných předpokladů pro  $T \rightarrow \infty$  v rámci konvergence v distribuci stejné rozdělení; rovněž pro  $T \rightarrow \infty$  platí

$$P(q \in CI^*) \rightarrow (1 - \alpha).$$

Navíc takto sestrojený interval  $CI^*$  pokrývá neznámý parametr s vyšší přesností, než intervaly konstruované na základě asymptotické normality statistiky  $R$ , viz také Prášková (2004).

Právě prezentovaný postup použijeme k analýze ekonometrického modelu životní pojišťovny (3.1.3) - (3.1.10) odhadnutého 3SLS-metodou (viz tabulka 3.2.1). Budeme porovnávat (asymptotické) 95% intervaly spolehlivosti využívající vlastností 3SLS-odhadu (viz odstavec 1.4.5), tj.

$$CI_{3SLS} = \left( \hat{l} - u_{0.975} \cdot \hat{s}(\hat{l}), \hat{l} + u_{0.975} \cdot \hat{s}(\hat{l}) \right),$$

kde  $\hat{l}$  označuje odhad parametru  $l$ ,  $\hat{s}(\hat{l})$  odhadnutou směrodatnou odchylku tohoto parametru a  $u_{0.975}$  kvantil normálního rozdělení, a 95% bootstrapové intervaly spolehlivosti, tj.

$$CI_{3SLS}^* = \left( \hat{l} - \gamma_{0.975}^* \cdot \hat{s}(\hat{l}), \hat{l} - \gamma_{0.025}^* \cdot \hat{s}(\hat{l}) \right),$$

kde příslušný  $p$ -kvantil  $\gamma_p^*$  získáváme na základě studentizované statistiky

$$\tilde{R}^* = \frac{\hat{l}^* - \hat{l}}{\hat{s}^*(\hat{l}^*)},$$

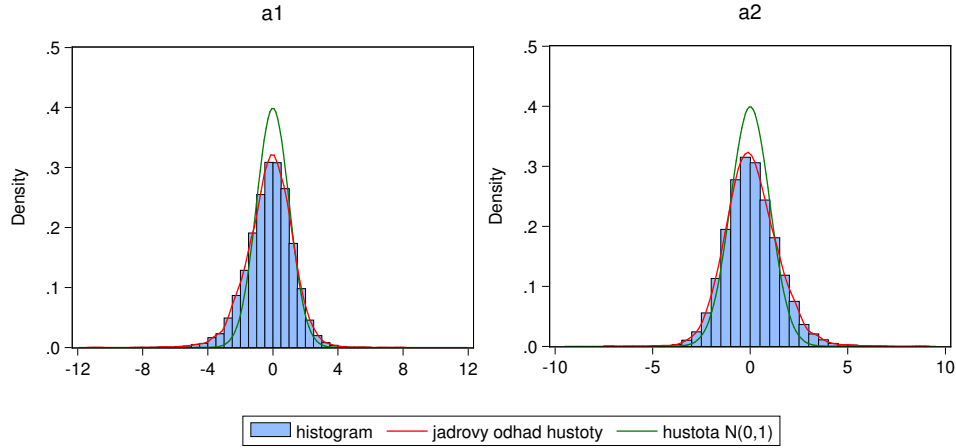
ve které odhadnutou směrodatnou odchylku  $\hat{s}^*(\hat{l}^*)$  počítáme standardními procedurami uvedenými v odstavci 1.4.5, podrobnosti nabízí Faure (2004, kapitola 9).

V tabulce 3.3.1 uvádíme srovnání hodnot výše uvedených intervalů  $CI_{3SLS}$  a  $CI_{3SLS}^*$  pro parametry našeho ekonometrického modelu (3.1.3) - (3.1.10) odhadnutého 3SLS-metodou. Praktický výpočet byl proveden pomocí procedury naprogramované v prostředí EViews 6.0 (tato je uvedena v příloze práce, viz strana 67), ve které jsme generovali 10 000 bootstrapových residuí.

Je zřetelné, že intervaly  $CI_{3SLS}^*$  jsou oproti  $CI_{3SLS}$  zpravidla širší, navíc, jak bylo uvedeno výše, pokrývají s vyšší přesností hodnoty neznámého parametru. Vzhledem k malému počtu sledovaných časových období je v tomto konkrétním případě použití (asymptotických) intervalů  $CI_{3SLS}$  značně diskutabilní. Tuto situaci dokumentuje jak poměrně zřetelný rozdíl šířek obou druhů intervalů, tak například obrázek 3.3.1, na němž je zobrazen histogram statistik  $\tilde{R}_1^*, \dots, \tilde{R}_{10000}^*$ , včetně jádrového odhadu hustoty, pro parametry  $a_1$  a  $a_2$  ve srovnání s hustotou normovaného normálního rozdělení (analogické grafy vycházejí i pro další parametry). Další numerické výsledky a komentáře lze najít v Faure (2004, kapitola 9).

	$CI_{3SLS}(D)$	$CI_{3SLS}(H)$	$CI_{3SLS}^*(D)$	$CI_{3SLS}^*(H)$
$a_1$	-76410.80	49678.29	-88074.82	87470.76
$a_2$	1.954231	2.644085	1.778534	2.718313
$b_1$	-19756.47	28586.87	-79766.50	37731.51
$b_2$	-0.386298	0.243129	-0.452694	2.086208
$b_3$	0.040153	0.109722	-0.203209	0.117446
$b_4$	-0.956819	7.371238	-5.299400	23.47777
$c_1$	-139502.2	285558.9	-206853.5	483485.0
$c_2$	-0.773509	0.941999	-2.502894	3.490478
$c_3$	-0.618457	2.409816	-2.178540	5.493698
$c_4$	-2.013834	20.01227	-33.74736	30.83075
$d_1$	-5704.730	869.7173	-7684.641	3819.961
$d_2$	0.192014	0.261023	0.039484	0.303262
$d_3$	-0.201091	0.300531	-0.495960	0.601540
$d_4$	-0.053935	-0.027034	-0.071419	0.045043
$e_1$	2285.693	3513.205	1826.923	4044.244
$e_2$	-0.001504	0.008305	-0.009722	0.013287
$e_3$	-0.407624	-0.307249	-0.514883	-0.236573
$e_4$	0.110236	0.126856	0.100696	0.145686
$f_1$	297.9674	1387.614	-209.0493	1881.502
$f_2$	0.001428	0.006477	-0.000933	0.008811
$f_3$	0.065780	0.130988	0.010888	0.195491
$f_4$	-0.220933	-0.053703	-0.386873	0.097812
$g_1$	-21362.08	111.9681	-28076.73	7970.249
$g_2$	-0.156880	0.391545	-0.478577	0.670632
$g_3$	0.148474	0.278828	0.084374	0.366039
$g_4$	-0.159508	0.415656	-0.501235	0.703436
$h_1$	-14085.80	1923.867	-18371.82	4803.168
$h_2$	0.847980	1.009562	0.818929	1.051391

**Tabulka 3.3.1:** Srovnání intervalů spolehlivosti  $CI_{3SLS}$  a  $CI_{3SLS}^*$  pro parametry ekonomického modelu (3.1.3) - (3.1.10) odhadnutého 3SLS-metodou, ( $D$ ), resp. ( $H$ ) označuje dolní, resp. horní mez příslušného intervalu. Generovali jsme 10 000 bootstrapových residuí.



**Obrázek 3.3.1:** Histogram statistik  $\tilde{R}_1^*, \dots, \tilde{R}_{10000}^*$ , včetně jádrového odhadu hustoty, pro parametry  $a_1$  a  $a_2$  ve srovnání s hustotou normovaného normálního rozdělení.

V následujícím odstavci 3.4 se budeme věnovat analýze scénářů budoucího vývoje, ve které můžeme opět využít metody bootstrap. Obecně se nabízejí dvě možnosti analýzy, a to pomocí deterministického přístupu a stochastické simulace (viz EViews 6.0).

Deterministický přístup postupuje v souladu s odstavcem 1.8, tj. na základě daných hodnot striktně exogenních proměnných, potřebných hodnot predeterminovaných proměnných a odhadů parametrů redukovaného tvaru dynamické soustavy simultánních rovnic (1.7.3), resp. (1.7.5) postupně podle (1.8.14) vypočteme  $\hat{\mathbf{y}}_1, \dots, \hat{\mathbf{y}}_{(T+\tau)}$ ,  $\tau \in \mathbb{N}$ .

Stochastická simulace pracuje analogicky, určitým způsobem však zohledňuje náhodnost residuální složky. Konkrétně na základě daných hodnot striktně exogenních proměnných, potřebných hodnot predeterminovaných proměnných a odhadů parametrů redukovaného tvaru dynamické soustavy simultánních rovnic (1.7.3) (podobně pak v (1.7.5)) získáme  $\hat{\mathbf{y}}_1, \dots, \hat{\mathbf{y}}_{(T+\tau)}$ ,  $\tau \in \mathbb{N}$ , jako

$$\hat{\mathbf{y}}_t = \mathbf{x}_t \hat{\mathbf{\Pi}} + \sum_{k=1}^p \hat{\mathbf{y}}_{(t-k)} \hat{\mathbf{\Delta}}_k + \tilde{\mathbf{v}}_t, \quad t = 1, \dots, T + \tau, \quad (3.3.6)$$

přičemž pokládáme  $\hat{\mathbf{y}}_{(-p+1)} = \mathbf{y}_{(-p+1)}, \dots, \hat{\mathbf{y}}_0 = \mathbf{y}_0$  a  $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_{(T+\tau)}$  jsou náhodně generované vektory.

Tento postup  $N$ -krát opakujeme ( $N \in \mathbb{N}$  volíme dostatečně velké), jednotlivé dílčí výsledky poté zprůměrujeme, a tak získáme definitivní výsledek stochastické simulace.

Náhodné vektory  $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_{(T+\tau)}$  můžeme postupně generovat:

- z mnohorozměrného normálního rozdělení s nulovou střední hodnotou a varianční maticí získanou při odhadu parametrů soustavy;
- metodou bootstrap, konkrétně při odhadu soustavy obdržíme  $T$  odhadnutých residuí (viz výše), z nichž poté náhodným výběrem s vracením dostaneme posloupnost  $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_{(T+\tau)}$ .

Srovnání deterministického přístupu a obou typů stochastických simulací provedeme v následujícím odstavci v rámci analýzy scénářů ekonometrického modelu životní pojišťovny.

### 3.4 Analýza scénářů

Jak již bylo zmíněno výše, v tomto odstavci se budeme věnovat analýze scénářů budoucího vývoje v rámci odhadnutého ekonometrického modelu životní pojišťovny (3.1.3) - (3.1.10), opět použijeme výsledky 3SLS-odhadu, viz tabulka 3.2.1. Konkrétně se budeme zajímat o vývoj vysvětlovaných proměnných modelu v letech 2006 až 2009. Ve třech uvažovaných scénářích mezi sebou porovnáme deterministický přístup a metody stochastické simulace, viz předchozí odstavce 3.3. Veškeré výpočty byly provedeny v programovém prostředí EViews 6.0. Ještě poznamenejme, že ve všech variantách výpočtů budeme uvažovat stabilní technickou úrokovou míru  $I_t$ ,  $t = 2006, \dots, 2009$ , ve výši 2,40 % p.a., tj. horní hranice pro technickou úrokovou míru v souladu s platnou regulací pojištnictví v ČR ve sledovaném období. Při stochastické simulaci budeme provádět 10 000 opakování.

#### První scénář

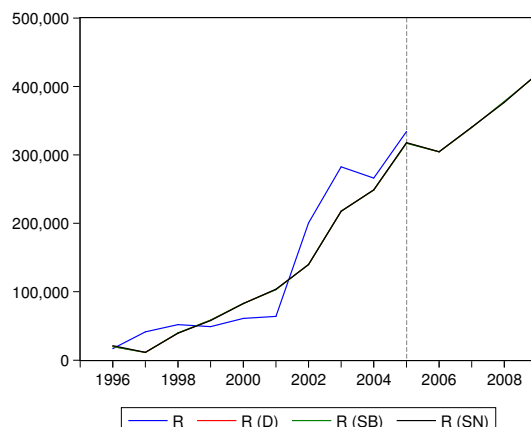
V tabulce 3.4.1 jsou obsažena relevantní data k prvnímu scénáři. Předpokládejme postupný nárůst počtu nových pojistných smluv, tj. průměrný počet pojistných smluv z období 1996 až 2005 (14 297 ks) každý rok zvyšujeme o 1 000 smluv, počet ukončených pojistných smluv v daném období volíme konstantní, jakožto průměrný počet ukončených smluv v období 1996 až 2005 (4 100 ks), eficientní úroková míra a koeficient růstu průměrné pojistné zůstávají konstantní.

Na obrázku 3.1.1 vidíme, že předpovědi i vyrovnané hodnoty proměnné  $R_t$  jsou ve všech třech sledovaných případech velmi blízké.

rok	2006	2007	2008	2009
$I_t^{(1)}$	2.40	2.40	2.40	2.40
$IEF_t^{(1)}$	6.00	6.00	6.00	6.00
$G_t^{(1)}$	2.50	2.50	2.50	2.50
$N_t^{(2)}$	14297	15297	16297	17297
$EX_t^{(2)}$	4100	4100	4100	4100
$R_t^{(3)} (D)$	304375	339981	377468	417418
$R_t^{(3)} (SN)$	304714	340405	376636	418100
$R_t^{(3)} (SB)$	304364	340046	377953	417434

<sup>(1)</sup> - % p.a., <sup>(2)</sup> - ks, <sup>(3)</sup> - tis. Kč.

**Tabulka 3.4.1:** Analýza prvního scénáře. Porovnááme mezi sebou výsledek životní pojišťovny  $R_t$  získaný deterministickým přístupem (D) a stochastickou simulací založenou na normálním rozdělení (SN) a bootstrapu (SB).



**Obrázek 3.4.1:** Analýza prvního scénáře. Průběh výsledku životní pojišťovny  $R_t$  získaného deterministickým přístupem (D) a stochastickou simulací založenou na normálním rozdělení (SN) a bootstrapu (SB).

### Druhý scénář

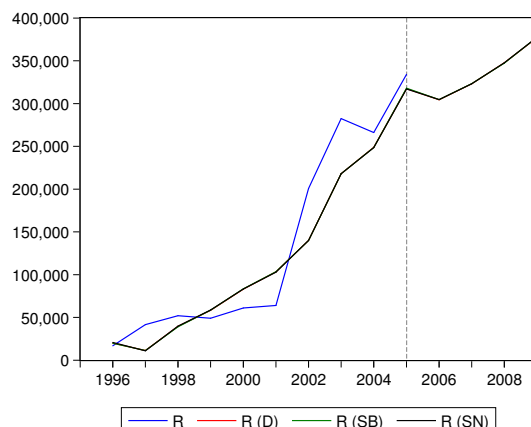
V tabulce 3.4.2 uvádíme data potřebná k analýze druhého scénáře, uvažujeme stejné hodnoty proměnných  $I_t$ ,  $IEF_t$  a  $G_t$  jako v prvním scénáři, dále předpokládáme stagnaci uzavírání nových pojistných smluv, tj. pro všechna období použijeme průměrný počet pojistných smluv z období 1996 až 2005 (14 297 ks), a nárůst počtu ukončených smluv, kdy se každý rok průměrný počet ukončených smluv v období 1996 až 2005 (4 100 ks) postupně zvýší o 500, 400, 300 ks.

Na obrázku 3.4.2 vidíme, že všechny tři studované přístupy poskytují velmi podobné předpovědi i vyrovnané hodnoty proměnné  $R_t$ , analogicky jako při analýze prvního scénáře.

rok	2006	2007	2008	2009
$I_t^{(1)}$	2.40	2.40	2.40	2.40
$IEF_t^{(1)}$	6.00	6.00	6.00	6.00
$G_t^{(1)}$	2.50	2.50	2.50	2.50
$N_t^{(2)}$	14297	14297	14297	14297
$EX_t^{(2)}$	4100	4600	5000	5300
$R_t^{(3)} (D)$	304375	323161	347727	377154
$R_t^{(3)} (SN)$	304545	323096	347574	377454
$R_t^{(3)} (SB)$	304616	323222	347348	377591

<sup>(1)</sup> - ‰ p.a., <sup>(2)</sup> - ks, <sup>(3)</sup> - tis. Kč.

**Tabulka 3.4.2:** Analýza druhého scénáře. Porovnáváme mezi sebou výsledek životní pojišťovny  $R_t$  získaný deterministickým přístupem (D) a stochastickou simulací založenou na normálním rozdělení (SN) a bootstrapu (SB).



**Obrázek 3.4.2:** Analýza druhého scénáře. Průběh výsledku životní pojišťovny  $R_t$  získaného deterministickým přístupem (D) a stochastickou simulací založenou na normálním rozdělení (SN) a bootstrapu (SB).

### Třetí scénář

Tabulka 3.4.3 obsahuje hodnoty proměnných třetího scénáře, v němž očekáváme postupné snižování počtu nově uzavřených pojistných smluv z průměrného počtu pojistných smluv z období 1996 až 2005 (14 297 ks) každoročně o 1 500 ks, dále uvažujeme snížení eficientní úrokové míry, resp. koeficientu růstu průměrné pojistné částky v období 2008 a 2009 na 5.00 % p.a., resp. 2.00 % p.a., ostatní hodnoty se shodují s druhým scénářem.

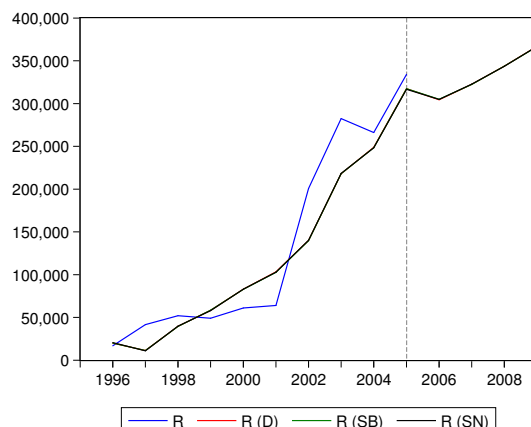
Na obrázku 3.4.3 vidíme, stejně jako v předchozích scénářích, že předpovědi i vyrovnané hodnoty proměnné  $R_t$  jsou ve všech třech případech velmi podobné.

rok	2006	2007	2008	2009
$I_t^{(1)}$	2.40	2.40	2.40	2.40
$IEF_t^{(1)}$	6.00	6.00	5.00	5.00
$G_t^{(1)}$	2.50	2.50	2.00	2.00
$N_t^{(2)}$	14297	12797	11297	9797
$EX_t^{(2)}$	4100	4600	5000	5300
$R_t^{(3)} (D)$	304375	322428	343720	366895
$R_t^{(3)} (SN)$	304756	322743	343544	366877
$R_t^{(3)} (SB)$	304926	322253	343765	366496

<sup>(1)</sup> - % p.a., <sup>(2)</sup> - ks, <sup>(3)</sup> - tis. Kč.

**Tabulka 3.4.3:** Analýza třetího scénáře. Porovnáváme mezi sebou výsledek životní pojišťovny  $R_t$  získaný deterministickým přístupem (D) a stochastickou simulací založenou na normálním rozdělení (SN) a bootstrapu (SB).



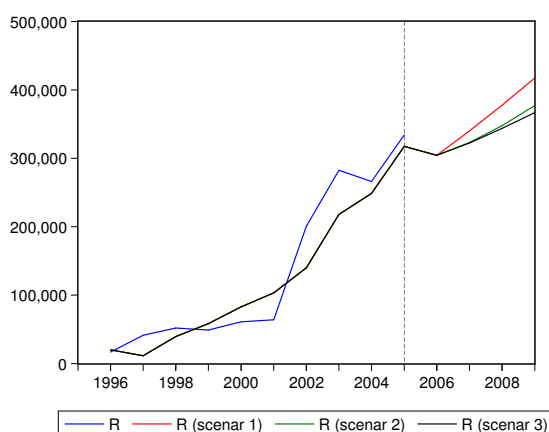


**Obrázek 3.4.3:** Analýza třetího scénáře. Průběh výsledku životní pojišťovny  $R_t$  získaného deterministickým přístupem (D) a stochastickou simulací založenou na normálním rozdělení (SN) a bootstrapu (SB).

### Shrnutí

Naznačili jsme možnosti analýzy scénářů budoucího vývoje v ekonometrickém modelu finančních toků v životní pojišťovně (3.1.3) - (3.1.10). Zaměřili jsme se především na zkoumání výsledku životní pojišťovny  $R_t$ , porovnávali jsme hodnoty získané deterministickým přístupem a stochastickou simulací, přičemž jsme dospěli k závěru, že při vysokém počtu opakování stochastických simulací (10 000) se jednotlivé výsledky příliš neliší. Pro upřesnění uvedme, že k výpočtu vyrovnaných či předpovědních hodnot  $R_t$  bylo samozřejmě nutné spočítat ostatní relevantní proměnné (viz (3.1.2)), ty zde však explicitně neuvádíme.

Na obrázku 3.4.4 je provedeno grafické srovnání jednotlivých scénářů na základě deterministického přístupu.



**Obrázek 3.4.4:** Grafické srovnání všech třech scénářů. Průběh jednotlivých výsledků životní pojišťovny  $R_t$  získaných deterministickým přístupem.

# Závěr

Cílem této práce byla hlubší prezentace teoretických východisek a praktická aplikace ekonometrických soustav (lineárních) simultánních rovnic.

V první kapitole práce jsme se zabývali především teoretickými aspekty tohoto ekonometrického konceptu. Zaměřili jsme se například na problematiku identifikace, nekonzistence OLS-odhadů, testování některých hypotéz, které úzce souvisejí s danou oblastí, dále na otázky dynamických soustav či konstrukcí předpovědí v rámci představených modelů. Značná pozornost byla věnována přehledu odhadových procedur a jejich vlastnostem.

Druhá kapitola shrnula základní pojmy životního pojištění potřebné pro následnou praktickou aplikaci.

Ve třetí kapitole textu byla prezentována vlastní implementace předchozích teoretických poznatků a výsledků v rámci ekonometrického modelu finančních toků v životní pojišťovně působící na českém pojistném trhu. K odhadu parametrů příslušné soustavy simultánních rovnic jsme zvolili dva nejběžnější přístupy, konkrétně 2SLS-odhad a 3SLS-odhad, na základě jejich vlastností představených v první kapitole jsme ze zřejmých důvodů pro další výpočty upřednostnili druhý jmenovaný. Dále jsme provedli testy některých vybraných hypotéz, například Hausmanův test specifikace, které dokumentovaly vhodnost uplatněného modelu. V neposlední řadě jsme uvedli metodu bootstrap a možnosti jejího použití v rámci studované problematiky. Představili jsme obecné schéma residuálního bootstrapu, který se v kontextu soustav simultánních rovnic obvykle používá. Srovnali jsme intervaly spolehlivosti pro jednotlivé parametry modelu získané pomocí normální aproximace a bootstrapu. Na závěr jsme analyzovali předpovědi, které jsme obdrželi na základě předem daných scénářů možného budoucího vývoje vybraných proměnných v odhadnutém ekonometrickém modelu. Naznačili jsme tak jednu z vysoce relevantních aplikací představeného modelu životní pojišťovny.

# Příloha A

## Bootstrapová procedura

V této příloze představíme programovou proceduru napsanou v EViews 6.0, s jejíž pomocí jsme mohli v odstavci 3.3 prakticky pracovat s residuálním bootstrapem. Poznamenejme, že příslušné vysvětlení veškerých použitých funkcí z prostředí EViews nalezneme v manuálu k tomuto softwaru.

Pracovní soubor `model.wf1` obsahuje data všech proměnných, které jsme představili na začátku třetí kapitoly, dále soustavu rovnic ekonometrických rovnic (3.1.3) - (3.1.10) ve struktuře `lfs` a celý model, tj. uvedenou soustavu včetně výpočetních vztahů pro upravené proměnné, ve struktuře `lfm`.

**Zdrojový kód.** Následuje vlastní zdrojový kód bootstrapové procedury, ten je uložen v souboru `boot.prg`.

```
'otevreni zdrojoveho workfilu (nastavit korektni cestu !!!)
wfopen C:\...\model.wf1

'priprava matice residui, centrovani
matrix MRES=RES 'RES obsahuje vsechna vypoctena residua

for !i=1 to @rows(MRES)
rowplace(MRES,@rowextract(MRES,!i)-@transpose(@cmean(MRES)),!i)
next

'priprava matice HGamma a IGamma
matrix HGamma=-@identity(@columns(mres))
HGamma(1,3)=lfs.@coefs(9)
HGamma(2,3)=lfs.@coefs(10)
HGamma(1,4)=lfs.@coefs(14)
HGamma(2,5)=lfs.@coefs(18)
HGamma(1,7)=lfs.@coefs(25)
HGamma(2,7)=lfs.@coefs(26)
HGamma(7,8)=lfs.@coefs(28)
matrix IGamma=@inverse(HGamma)

'priprava matic vysledku
!N=10000

matrix(lfs.@ncoefs,!N) BCOEFS
matrix(lfs.@ncoefs,!N) BSTDS
```

```

'BOOTSTRAP
for !i=1 to !N
matrix RSMPL=@resample(MRES)
matrix BOOT=-RSMPL*IGamma

for !j=1 to @rows(MRES)
  lfm.solve(s=d,d=d)

  p(!j+1)=p_0(!j+1)+BOOT(!j,1)
  cs(!j+1)=cs_0(!j+1)+BOOT(!j,2)
  v(!j+1)=v_0(!j+1)+BOOT(!j,3)
  eac(!j+1)=eac_0(!j+1)+BOOT(!j,4)
  ead(!j+1)=ead_0(!j+1)+BOOT(!j,5)
  re(!j+1)=re_0(!j+1)+BOOT(!j,6)
  prof(!j+1)=prof_0(!j+1)+BOOT(!j,7)
  profs(!j+1)=profs_0(!j+1)+BOOT(!j,8)
  vi(!j+1) = (1 + i(!j+1)) * (v(!j+1) + v(!j)) / 2
  vief(!j+1) = (1 + ief(!j+1)) * (v(!j+1) + v(!j)) / 2
next

lfs.3sls

colplace(BCOEFs,lfs.@coefs,!i)
colplace(BSTDS,lfs.@stderrs,!i)

'radly naplnime puvodnimi daty, puvodni odhad
p=p_orig
cs=cs_orig
v=v_orig
eac=eac_orig
ead=ead_orig
re=re_orig
prof=prof_orig
profs=profs_orig
vi=vi_orig
vief=vief_orig

lfs.3sls
next

'po ukonceni bootstrapingu - puvodni data, puvodni odhad
p=p_orig
cs=cs_orig
v=v_orig
eac=eac_orig
ead=ead_orig
re=re_orig
prof=prof_orig
profs=profs_orig
vi=vi_orig
vief=vief_orig

lfs.3sls

```

```

matrix(@rows(BCOEFS),@columns(BCOEFS)) BR

for !i=1 to @rows(BCOEFS)
for !j=1 to @columns(BCOEFS)
    BR(!i,!j)=(lfs.@coefs(!i)-BCOEFS(!i,!j))/BSTDS(!i,!j)
next
next

'vypocet intervalu spolehlivosti
!a=0.05

if @floor(!N*!a)>=1 then
matrix(@rows(BCOEFS),4) IS

    for !i=1 to @rows(BCOEFS)
    IS(!i,1)=lfs.@coefs(!i)-@qnorm(1-!a/2)*lfs.@stderrs(!i)
    IS(!i,2)=lfs.@coefs(!i)-@qnorm(!a/2)*lfs.@stderrs(!i)
    IS(!i,3)=lfs.@coefs(!i)-
        @sort(@rowextract(BR,!i))(@floor((1-!a/2)*!N))*lfs.@stderrs(!i)
    IS(!i,4)=lfs.@coefs(!i)-
        @sort(@rowextract(BR,!i))(@floor(!a/2)*!N))*lfs.@stderrs(!i)
    next

endif

'matice pro vykresleni grafu
matrix graf=@transpose(BR)

'odstraneni nepotrebnych struktur
delete BOOT
delete RSMPL
delete MRES
delete HGamma
delete IGamma
delete *_0

```

Na závěr uvedme, že oba soubory model.wf1 a boot.prg jsou obsaženy na příloženém datovém médiu, které je nedílnou součástí práce.

# Literatura

- Anděl, J. (2007). *Základy matematické statistiky*. Praha: MatfyzPress.
- Baranoff, E. G., Papadopoulos, S., & Sager, T. W. (2007). Capital and risk revisited: A structural equation model approach for life insurers. *The Journal of Risk and Insurance*, 74(3), 653–681.
- Cipra, T. (1984). *Ekonometrie*. Praha: SPN.
- Cipra, T. (1998). Ekonometrická analýza finančních toků v životní pojišťovně. *Pojistné rozpravy*, 1998(3), 58–72.
- Cipra, T. (2006a). *Finanční a pojistné vzorce*. Praha: GRADA.
- Cipra, T. (2006b). *Pojistná matematika - teorie a praxe*. Praha: Ekopress.
- Cipra, T. (2008). *Finanční ekonometrie*. Praha: Ekopress.
- Davidson, R. & MacKinnon, J. G. (2004). *Econometric theory and methods*. New York: Oxford University Press.
- Dhrymes, P. J. (1994). *Topics in Advanced Econometrics: Linear and nonlinear simultaneous equations*. New York: Springer.
- Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7(1), 1–26.
- Faire, R. C. (2004). *Estimating How the Macroeconomy Works*. Cambridge, MA: Harvard University Press.
- Freedman, D. A. (1984). On bootstrapping two-stage least-squares estimates in stationary linear models. *The Annals of Statistics*, 12(3), 827–842.
- Greene, W. H. (2003). *Econometric Analysis*. New Jersey: Prentice Hall.
- Hall, A. R. (2005). *Generalized method of moments*. New York: Oxford University Press.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. New York: Springer.
- Prášková, Z. (2004). Metoda bootstrap. In Antoch, J. & Dohnal, G. (Eds.), *Sborník ROBUST 2004*, (pp. 299–314), Praha. JČMF.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*.  
Cambridge: MIT Press.