

STRUKTŪRINĖS LYGTYS (SEM)

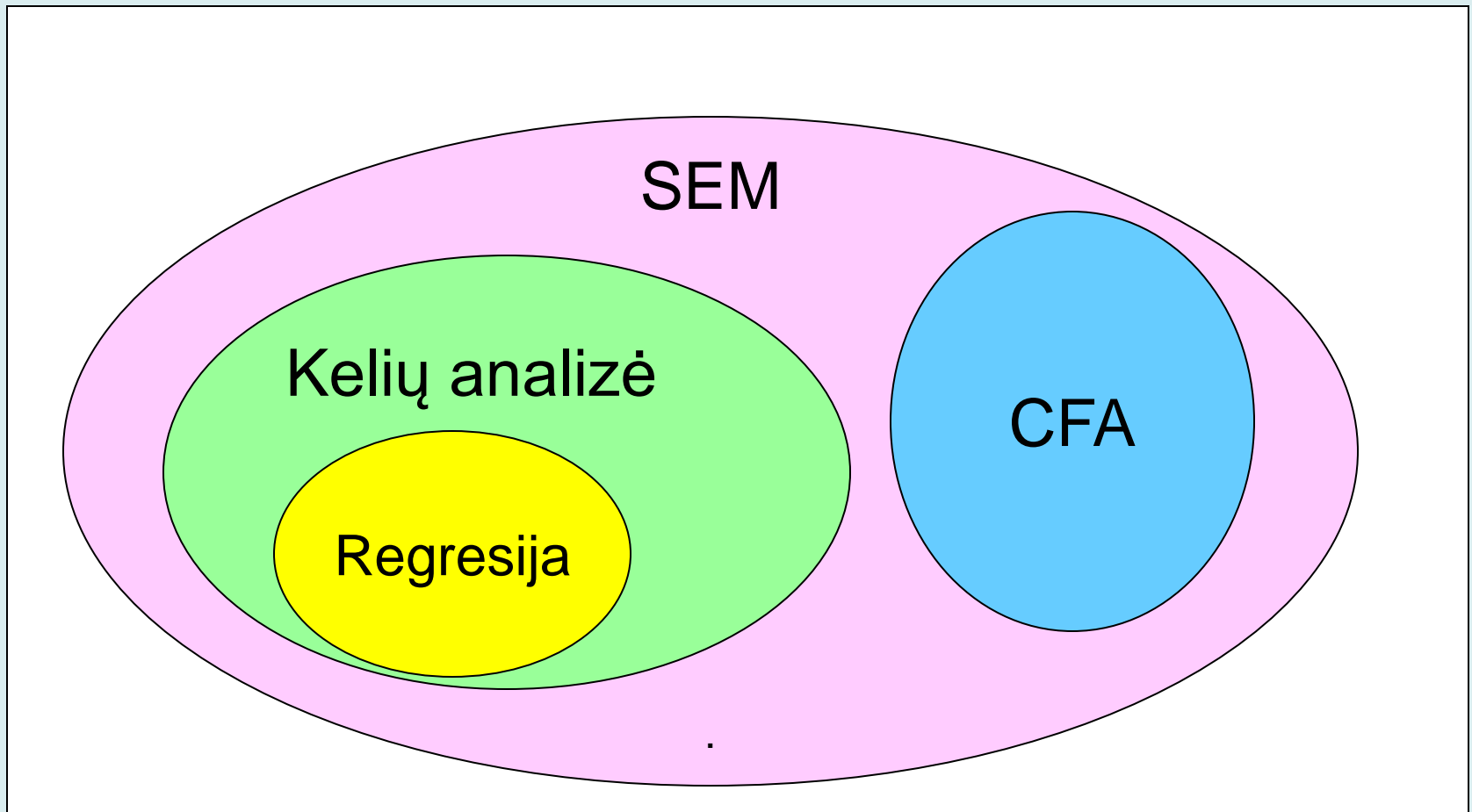
Ką daro

SEM analizė tikrina, ar tyrėjo sudarytas kintamųjų priklausomybės modelis tinka duomenims.

Gali būti ir latentinių (tiesiogiai nematuojamų) kintamųjų.

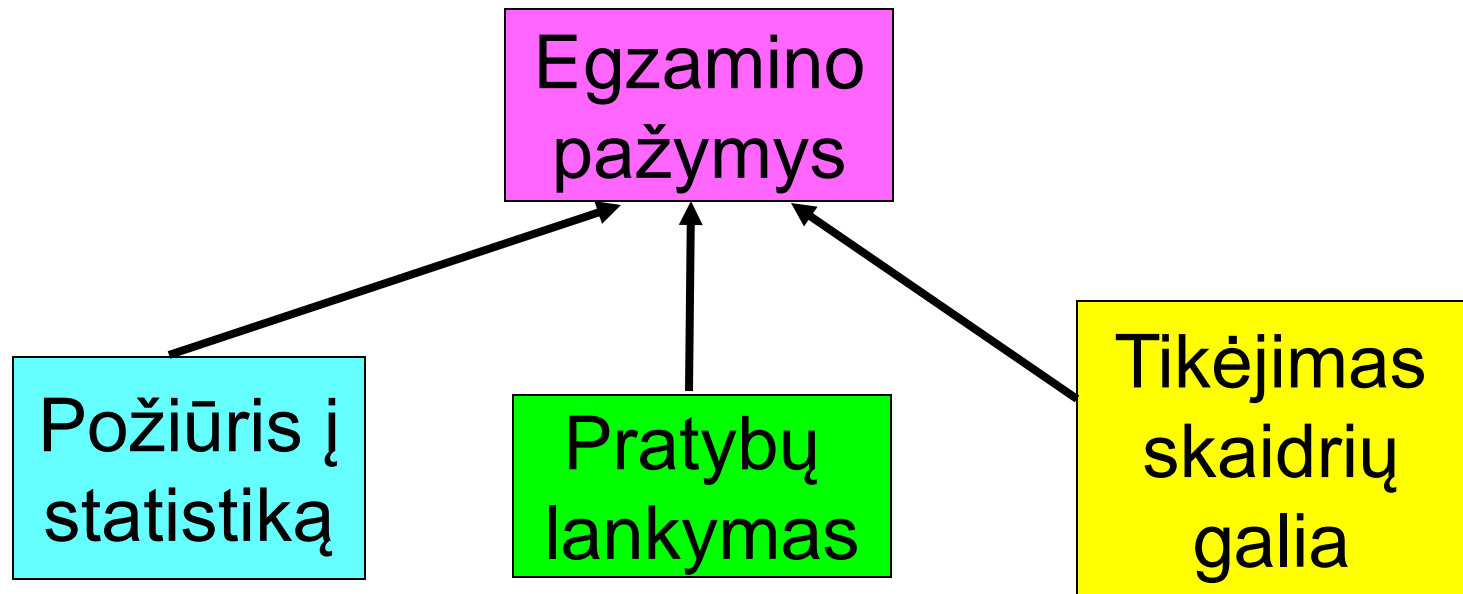
Skirtingi mokslai kūrė skirtingus metodus, o SEM juos sujungė į vieningą visumą.

Ryšiai su kitais metodais

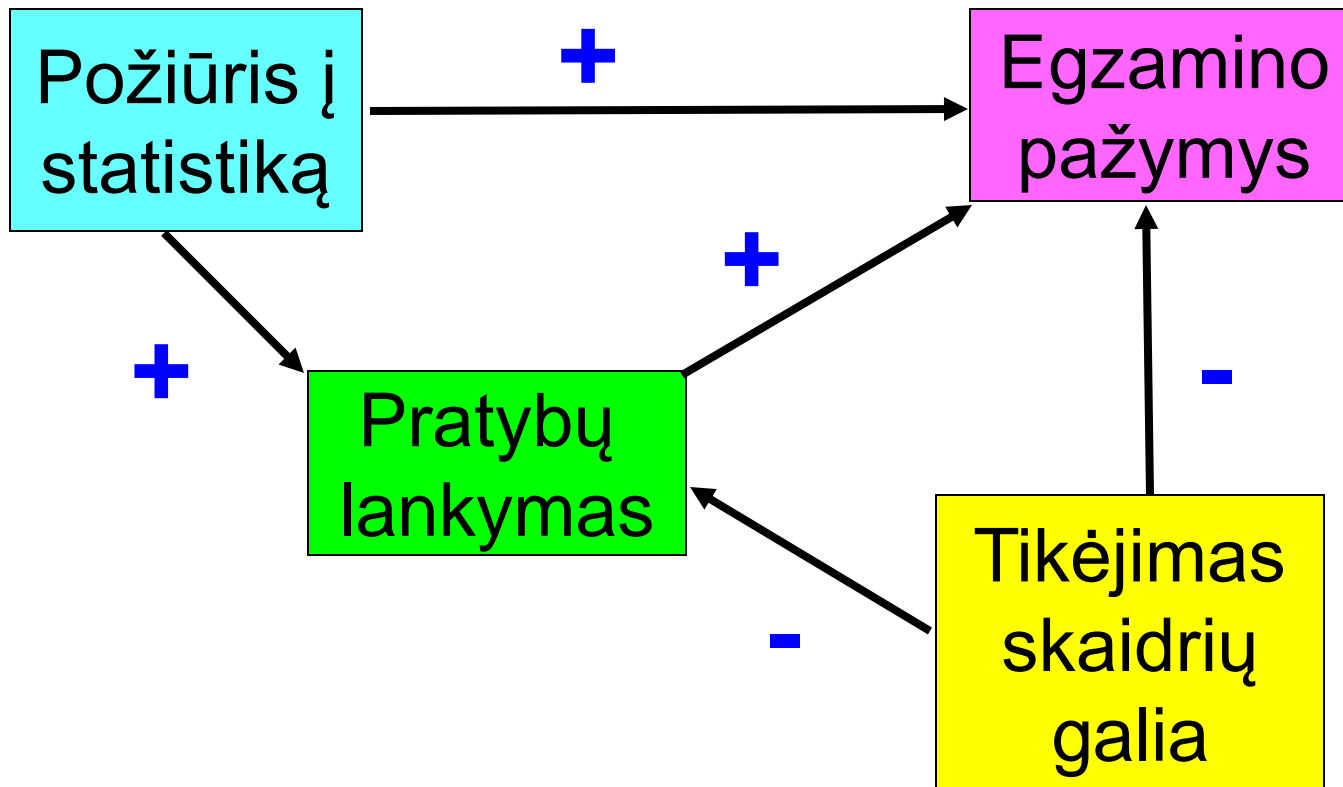


Regresijos idėja

Vieną intervalinį kintamąjį paaiškina kiti.
Visi kintamieji – stebimieji.



Kelių analizės (PA) idėja



CFA idėja

Stebimuosius kintamuosius sieja latentinio faktoriaus įtaka.

Žiūri LTV

Verda cepelinus

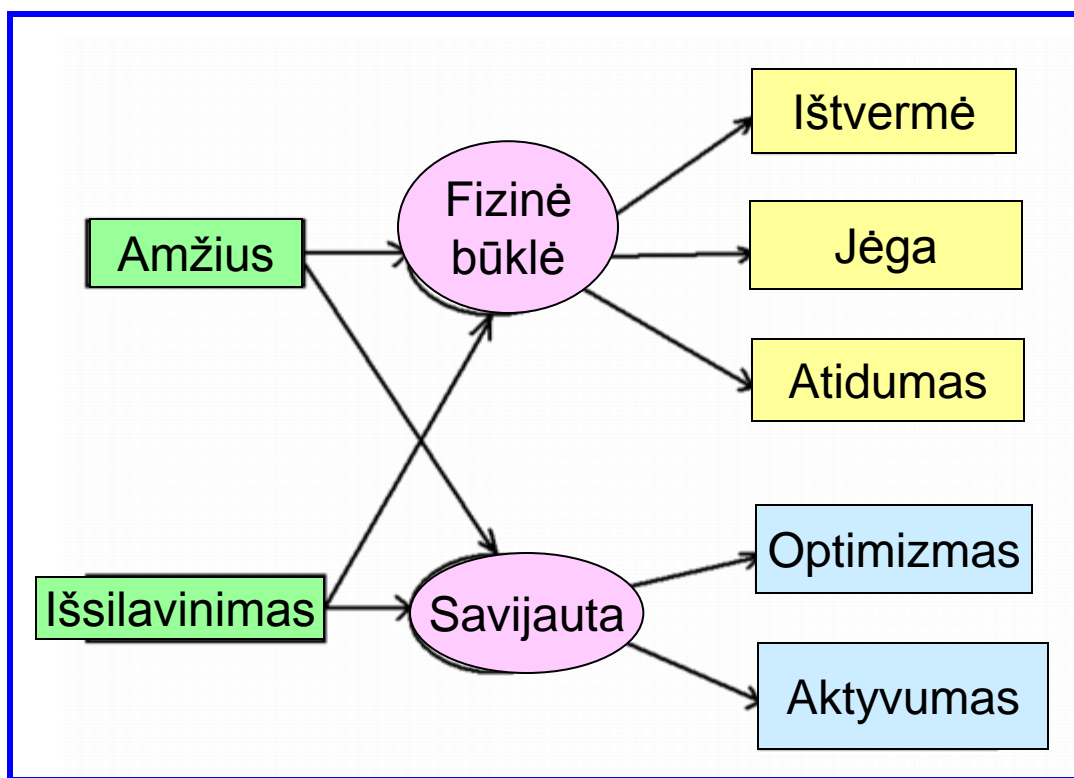
Pasiilgsta troleibusų

Nostalgija

```
graph TD; Nostalgija((Nostalgija)) --> Ziuri[Žiūri LTV]; Nostalgija --> Verda[Verda cepelinus]; Nostalgija --> Pasiilgsta[Pasiilgsta troleibusų];
```

Bendresnio SEM idėja

Susieja CFA su PA.



Tikslai

- Norime sužinoti, ar modelis apskritai tinka duomenims.
- Siekiame įvertinti kiekvieno kintamojo poveikio stiprumą (svorį, regresijos koeficientą).
- Norime įvertinti kintamųjų priklausomybes.

Pradiniai duomenys

Žinome matuojamų kintamųjų vidurkius, dispersijas ir koreliacijas arba (dažniau) **kintamųjų kovariacijų matricą**, t.y. visas įmanomas kovariacijas ir dispersijas.

Kovariacija matuoja kintamųjų priklausomybių stiprumą.

Pačių kintamųjų reikšmės nėra būtinos.

Kaip tikrinamas modelis

- Atkuriamas kintamųjų kovariacijų matrica, atsižvelgiant į spėjamas kintamųjų priklausomybes.
- Ši matrica lyginama su tikraja kovariacijų matrica.
- Jei mažai skiriasi (tuos skirtumus matuoja chi-kvadrato statistikos reikšmė), darome išvadą, kad duomenys spėjamam modeliui neprieštarauja.

Pastabos

- Gali būti keli (ir net labai skirtingi) modeliai, kurie vienodai gerai tinka duomenims.
- SEM palengvina tiesos paieškas, bet nepakeičia tyrėjo.
- Todėl išvadas teks daryti patiems, o po to įtikinti savo tiesa kitus.

Bendroji tyrimo schema

1. Modelio aprašymas.
2. Modelio identifikavimas.
3. Parametrų įverčių skaičiavimas.
4. Tikrinimas, ar tinka duomenims.
5. Modelio tobulinimas.

Modelio aprašymas

Tyrėjas sukuria savo spėjimą modelį. Tai darydamas jis:

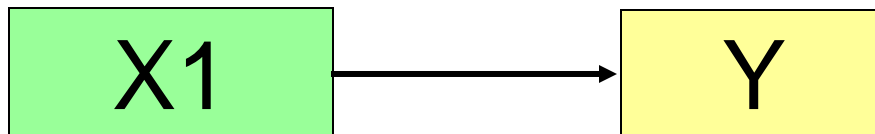
- remiasi savo mokslais;
- nebando sukišti į modelį kuo daugiau priklausomybių (o gal išeis);
- naudoja intervalinius (normaliuosius) kintamuosius.

Modelio aprašymas

Kintamieji, kurie paaiškinami kitais kintamaisiais (diagramoje turi ateinančią rodyklę), vadinami *endogeniniais*.

Kintamieji, kurie tokios rodyklės neturi, yra *egzogeniniai*.

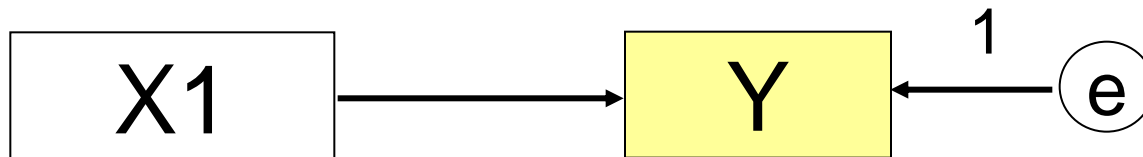
X1 egzogeninis, Y endogeninis.



Modelio aprašymas

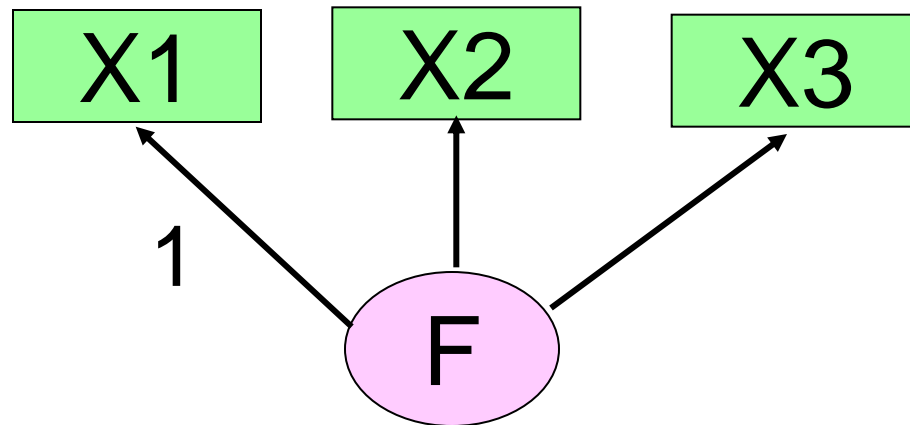
Jei manome, kad endogeninis kintamasis priklauso ne tik nuo modelio kintamųjų, tai AMOS nurodome, kad jis turi liekamąją paklaidą. Tos paklaidos dispersija taip pat bus įvertinta.

Kelias nuo paklaidos (paklaidos regresijos koeficientas) turi vienetinį svorį.



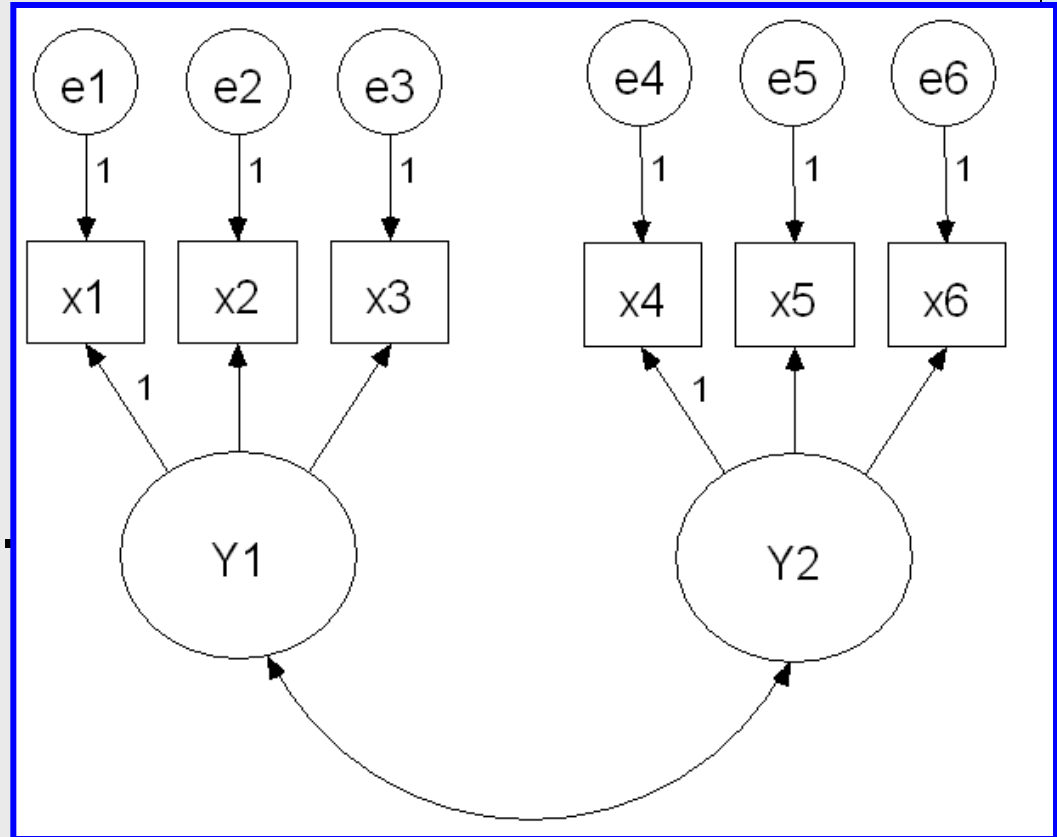
Modelio aprašymas

CFA atveju vienam iš kintamųjų priskiriamas vienetinis regresijos koeficientas (alternatyva – vieną dispersiją prilyginti vienetui).



Modelio aprašymas

Visi kintamieji
(ir latentiniai)
gali koreliuoti.
Koreliacija taip
pat bus įvertinta.



Modelio identifikavimas

Visus parametrus teks vertinti naudojantis kovariacijos matrica.

O joje yra tik $k(k+1)/2$ skirtingų kovariacijų (informacinių vienetų).

Todėl negalima į modelį įtraukti per daug parametrų.

Parametrai: regresijos svoriai, koreliacijos ir nežinomos dispersijos.

Modelio identifikavimas

Modeliai skirstomi į:

- neidentifikuojamus (*underidentified*) modelius (parametrų daugiau nei $k(k+1)/2$);
- tiksliai identifikuojamus (*just identified*) modelius (parametrų yra $k(k+1)/2$);
- nevienareikšmiškai identifikuojamus (*overidentified*) modelius (parametrų yra mažiau nei $k(k+1)/2$).

Modelio identifikavimas

Žinoma, galima sugalvoti visokių keistų priklausomybių, tačiau nesąmoningų schemų nupiešti AMOS neleis.

Patiems nereikia skaičiuoti vertinamų parametru ir modelio laisvės laipsnių. Jeigu modelis bus neidentifikuojamas, AMOS programa apie tai praneš.

Ką tikriname

- Ar bendrai modelis tinka duomenims.
- Kiekvieno parametro didumą ir prasmę.

Parametrų ženklai turi atitikti spėjama priklausomybę. Pavyzdžiui, bus blogai, jei pažymys neigiamai koreliuos su mokymosi laiku.

Bendrieji tikimo rodikliai

- **Chi kvadrato** kriterijus. Gerai, kai p reikšmė didesnė už 0,05.
- **GFI** (*goodnes of fit index*). Modelis gerai tinka duomenims, kai **GFI** arti 0,95.
- **RMSEA** (*root mean square error of approximation*). Modelis gerai tinka duomenims, kai **RMSEA** < 0,05.

Bendrieji tikimo rodikliai

- *NFI* (*normed fit index*). Modelis gerai tinka duomenims, kai *NFI* arti 0,95.
- *TLI* (*Tucker-Lewis index*). Modelis gerai tinka duomenims, kai *TLI* arti 0,95.
- *AIC* (Akaikės indeksas) naudojamas modelių palyginimui. Kuo *AIC* mažesnis, tuo modelis geresnis.

Papildoma informacija

Dažnai pateikiama informacija apie pilnąjį (*saturated*) modelį, t.y. modelį su visais įmanomais keliais, ir apie nepriklausomą (*independence*) modelį, kuriame nėra jokių kelių.

Hipotezės apie parametrus

Įvertinama, ar kiekvienas parametras statistiškai reikšmingai skiriasi nuo 0.

Regression Weights: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
INCOME <--- SAT	,011	,016	,690	,490	
INCOME <--- EDUCATN	3,166	2,965	1,068	,286	

Pavyzdžiui, čia abudu keliai abejotini nes p reikšmės didesnės už 0,05.

Modelio tobulinimas

Sudarome naują modelį, kurį vėl tiriamo.

AMOS galima iškart tirti keletą alternatyvių modelių.

AMOS galima palyginti keletą modelių.

Pavyzdžiui, pažiūrėti, ar berniukų ir mergaičių kelių koeficientai tokie pat.

AMOS

Yra laisvai prieinama studentiška **AMOS5** versija. Joje yra apribojimų kintamųjų skaičiui (iki 7) ir parametru skaičiui.

Čia naudota **SPSS AMOS16** versija.

Internete yra daug išsamių, mokymui skirtų, instrukcijų. Todėl susipažinsime tik su programos naudojimo pagrindais.

Pavyzdys

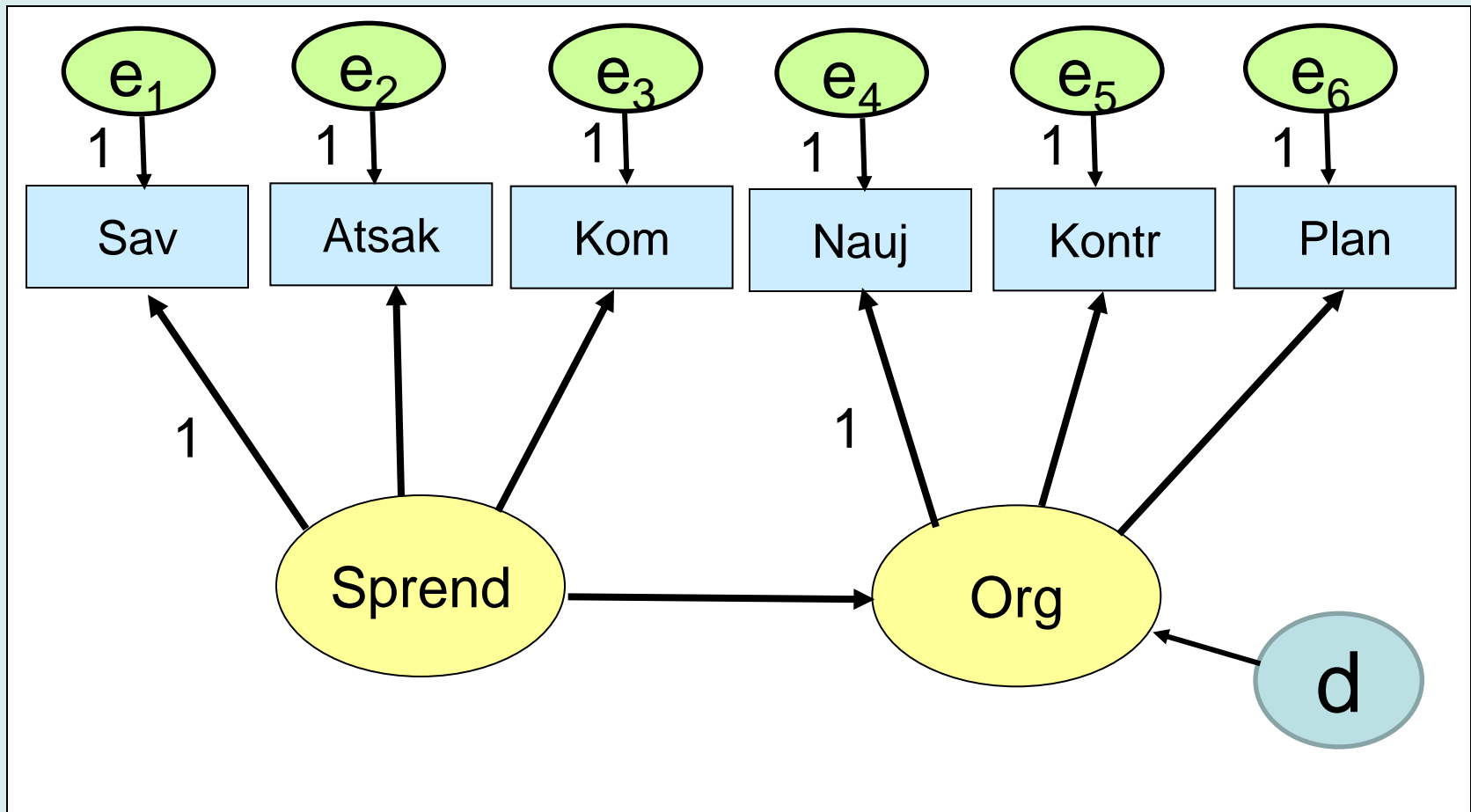
Spėjame, kad vadybininko profesinei veiklai įtaką daro du latentiniai faktoriai: organizaciniai gabumai (*Org*), pasireiškiantys darbo planavimu (*Plan*), kontroliavimu (*Kontr*) ir naujovių diegimu (*Nauj*); gebėjimas priimti sprendimus (*Sprend*), pasireiškiantis vadybininko savarankiškumu (*Sav*), atsakomybės jausmu (*Atsak*) ir mokėjimu bendrauti su kitais vadybininkais (*Kom*).

Modelio parinkimas

Patikrinsime SEM modelį, sudarytą knygos *Statistika ir jos taikymai III*, 3.3.1 pavyzdyje.

Iš pradžių nusibraižome spėjamų priklausomybių diagramą. Primename, kad visi objektai, į kuriuos ateina rodyklė, turi liekamąsiais paklaidas. Visos paklaidos ir vienas aiškinamasis kintamasis turi vienetinius svorius.

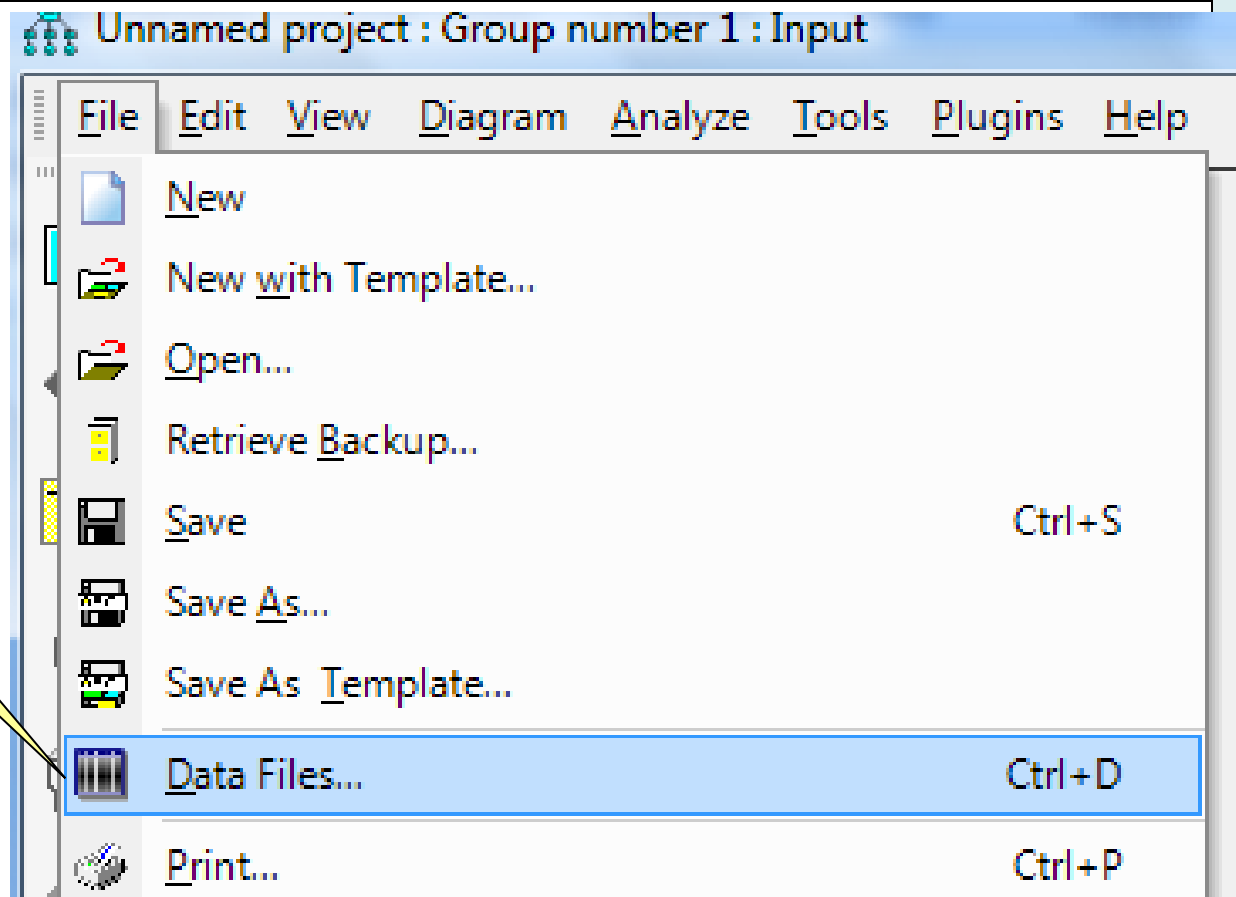
Modelio diagrama



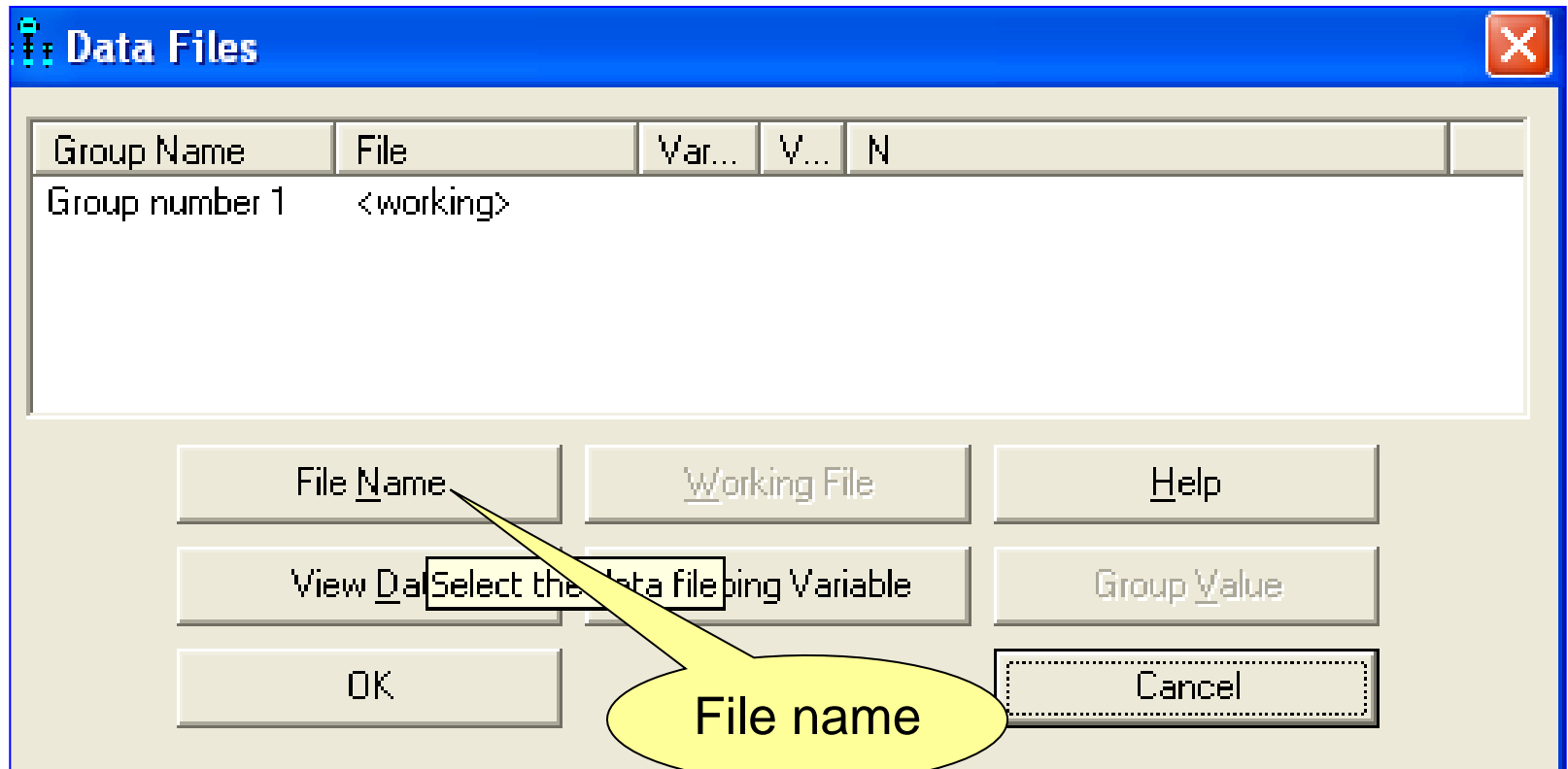
Failo atvėrimas

File →
Data files

Čia

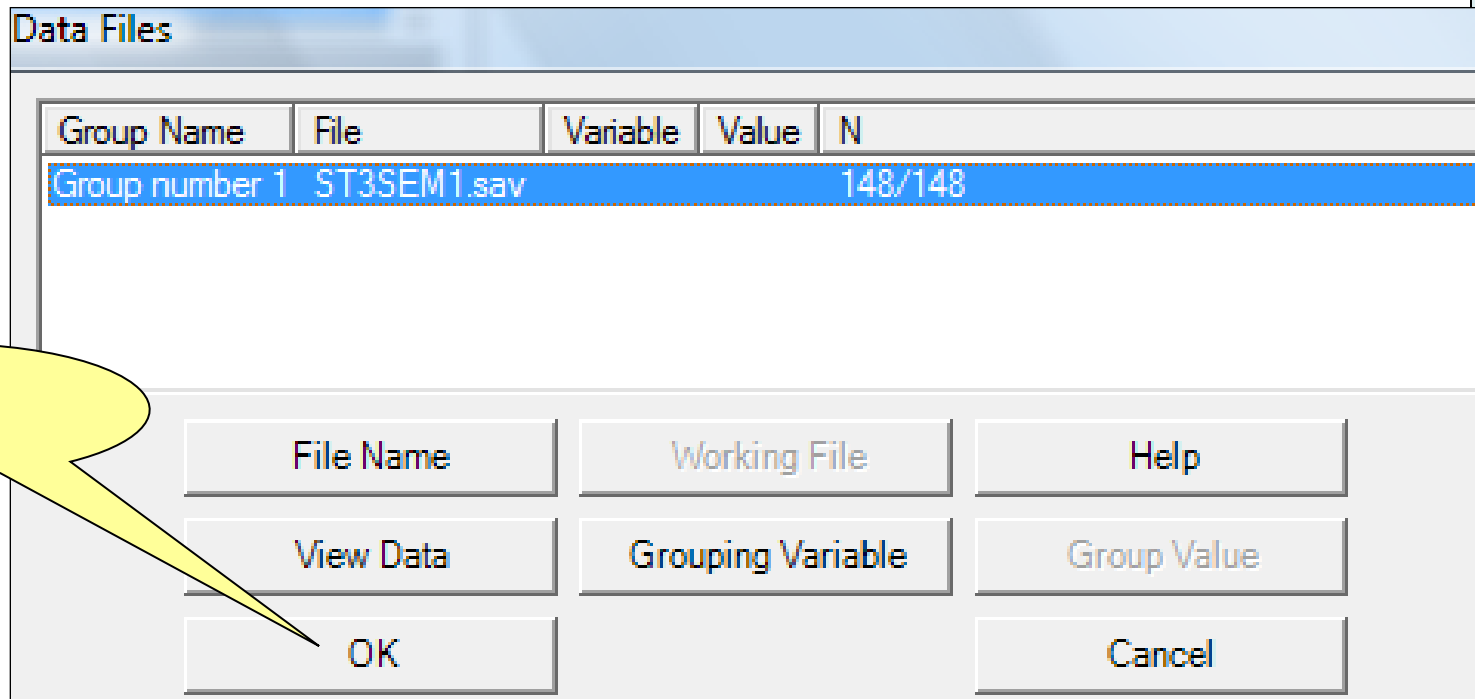


Failo atidarymas (atvėrimas)

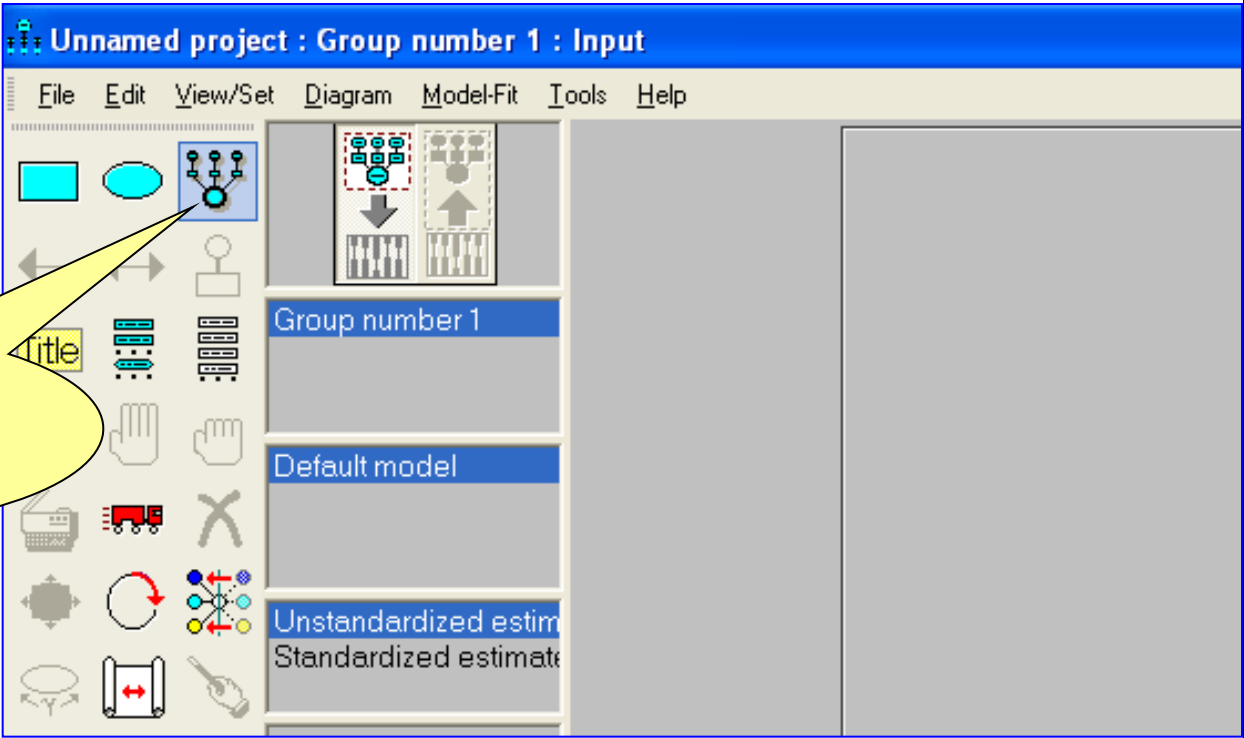


Failo atidarymas

Pasirenkame failą ST3SEM1.



Diagramos braižymas

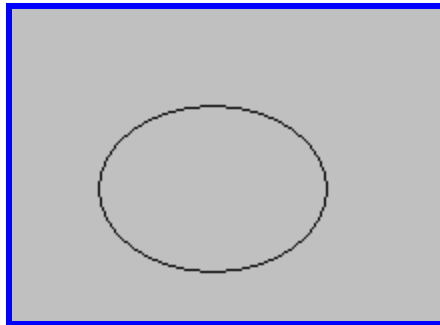


The screenshot shows a software window titled "Unnamed project : Group number 1 : Input". The menu bar includes "File", "Edit", "View/Set", "Diagram", "Model-Fit", "Tools", and "Help". A toolbar on the left contains various icons, with a tree diagram icon highlighted. A yellow callout bubble with the text "Paspaudžiame" (We click) points to this icon. Below the callout, a separate box shows a larger version of the tree diagram icon. The main workspace on the right contains a diagram with a central node and three branches, and a list of elements including "Group number 1", "Default model", "Unstandardized estim", and "Standardized estimate".

Paspaudžiame

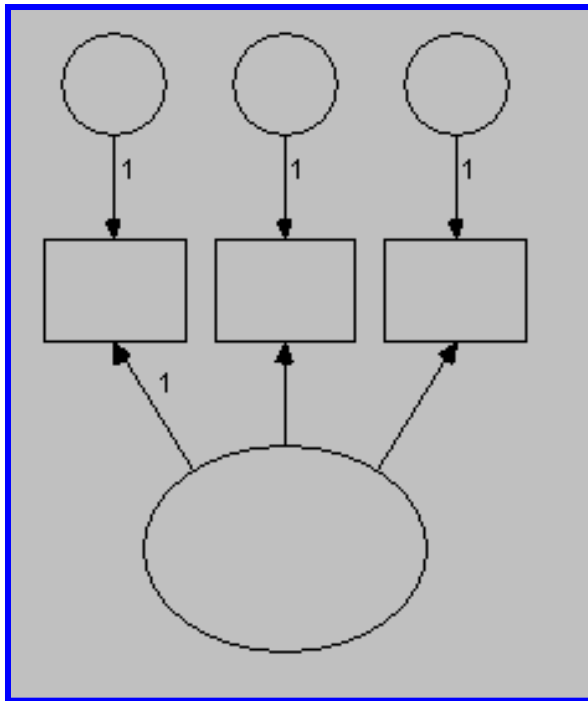
Diagramos braižymas

Nuspaudę kairinį klavišą, dešiniame lauke nupiešiame elipsę.



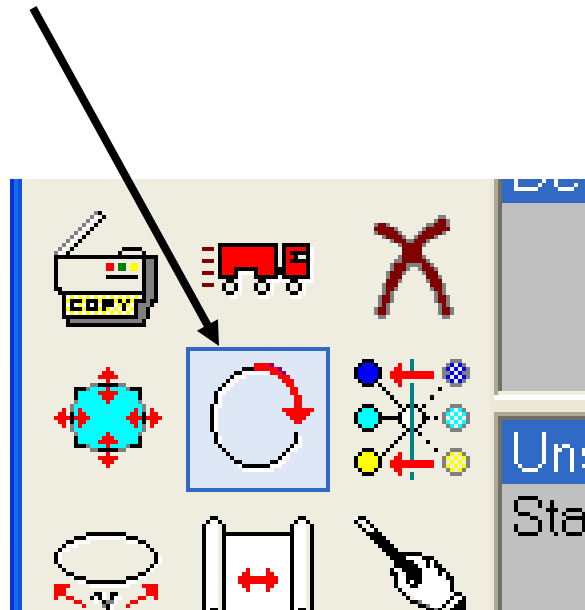
Diagramos braižymas

Triskart ant jos „spragtelime“ pele.



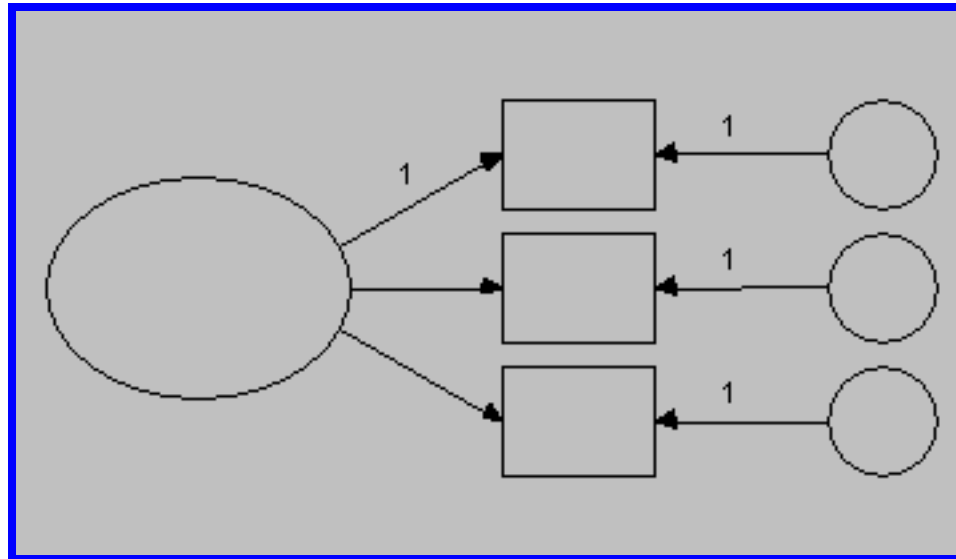
Diagramos braižymas

Spragtelime



Diagramos braižymas

Grįžtame prie piešinio ir „spregtelime“ pagrindinės elipsės viduje. Piešinys pasisuka.



Diagramos braižymas

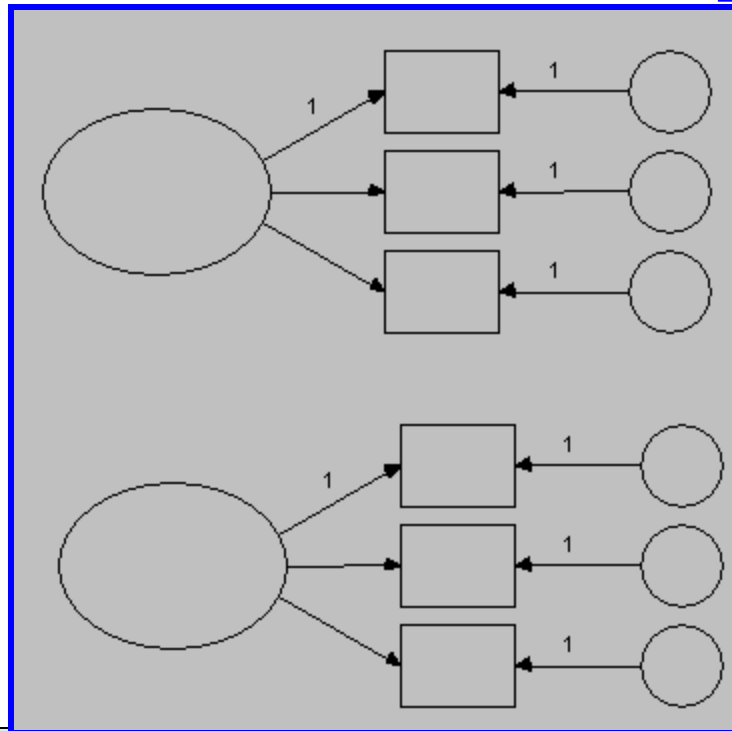
Pasirenkame  (*select all*),

Po to spragtelime  (ruošiamės kopijuoti).

Tada piešinyje (kuris pamėlynavo) nuspaudę kairijį klavišą, patempiame vidurinę kvadratą žemyn.

Diagramos braižymas

Pažymėjimus panaikiname su
Gauname

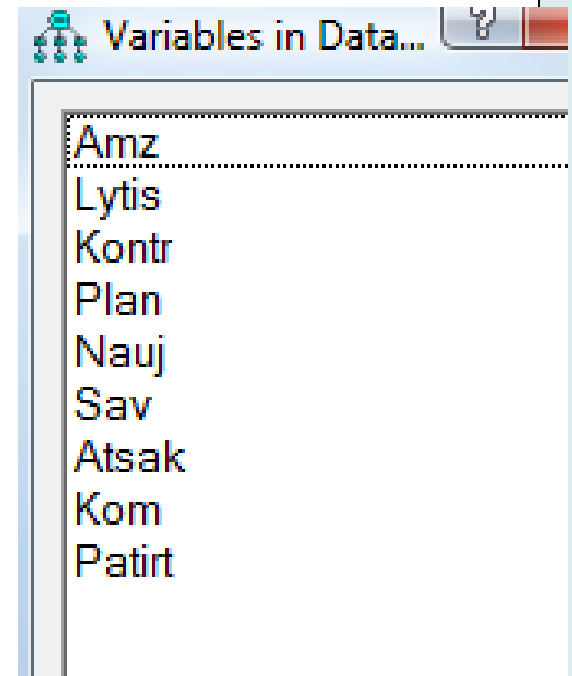


Diagramos braižymas

Spragtelime



Gauname kintamųjų sąrašą.



Kintamųjų vardus perkeliame
ant kvadratėlių. Užveriamo sąrašą.

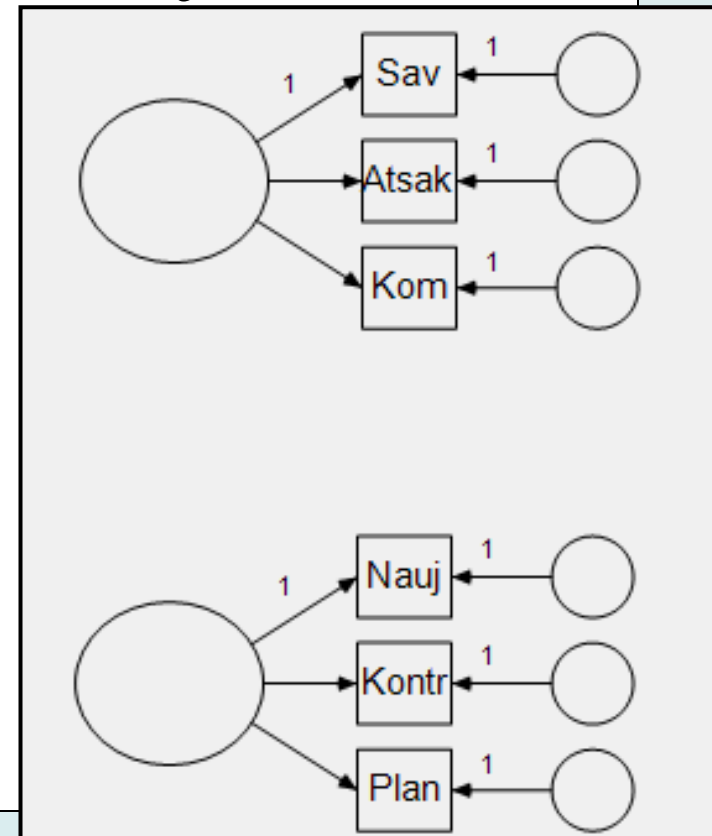
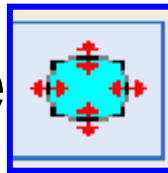
Diagramos braižymas

Gauname ne itin malonų vaizdą:

Viską pažymime
spragtelėdami



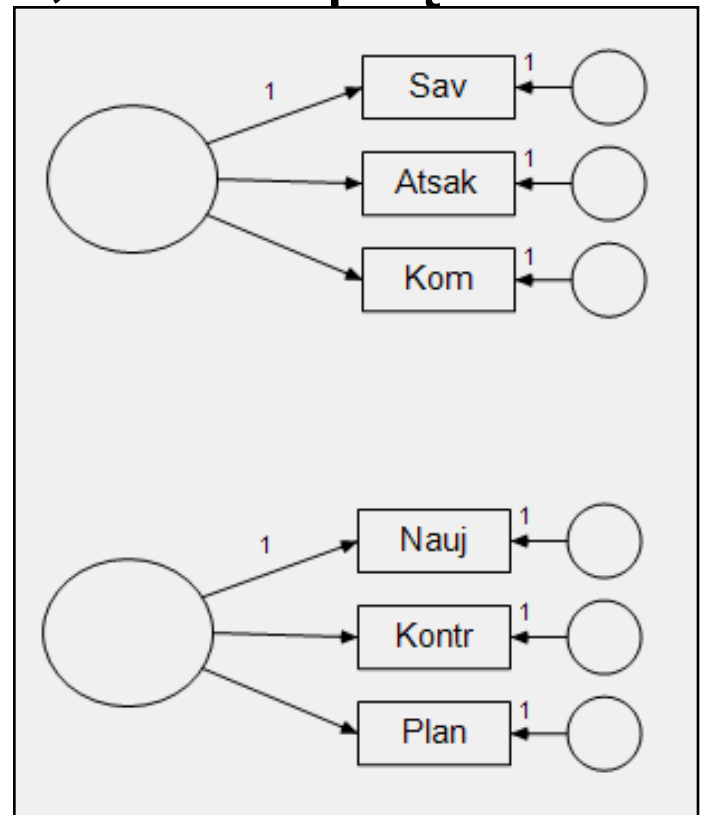
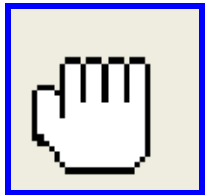
Tada spragtelime



Diagramos braižymas

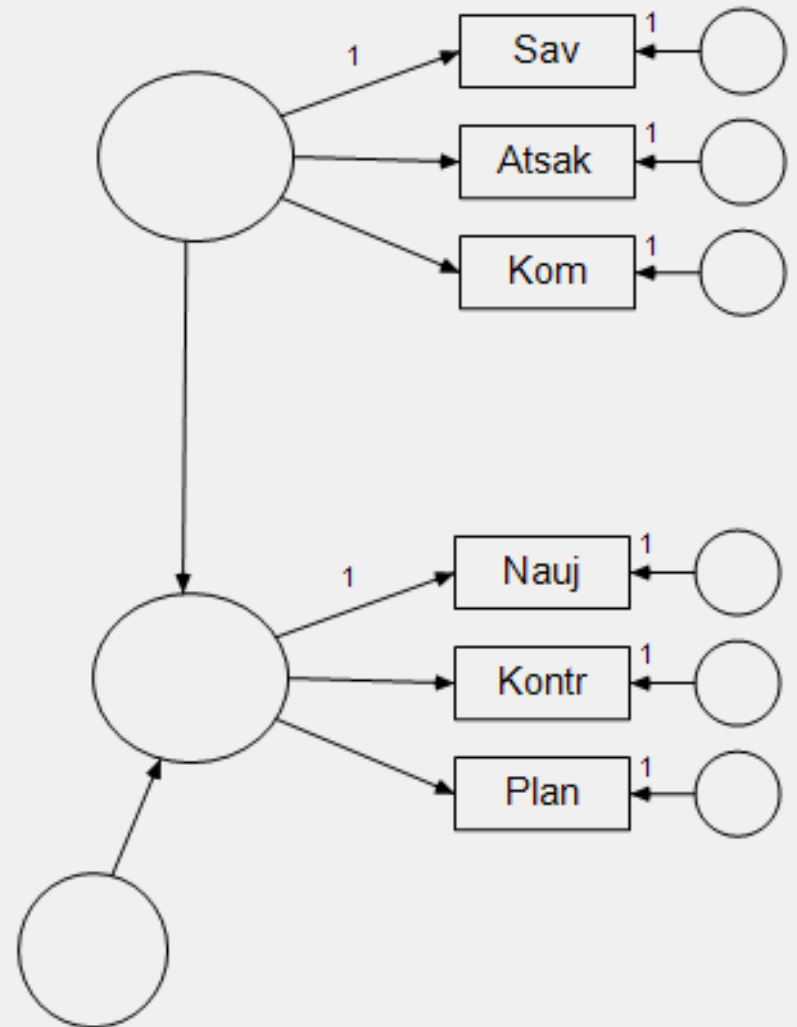
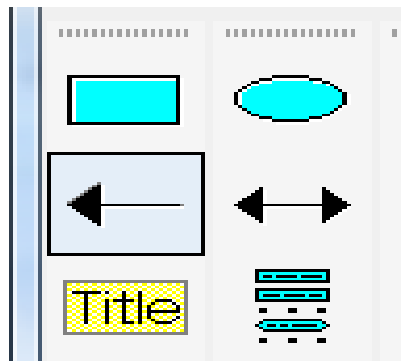
Patempiname kvadratėlius, kad taptų
stačiakampiais.

Po to spragtelime



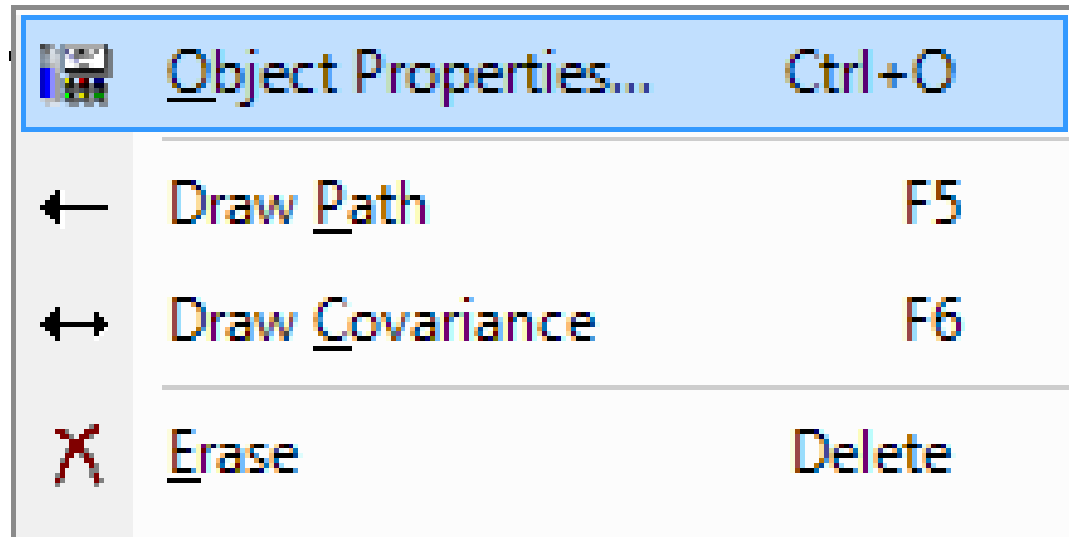
Diagramos braižymas

Papildome brėžinį.



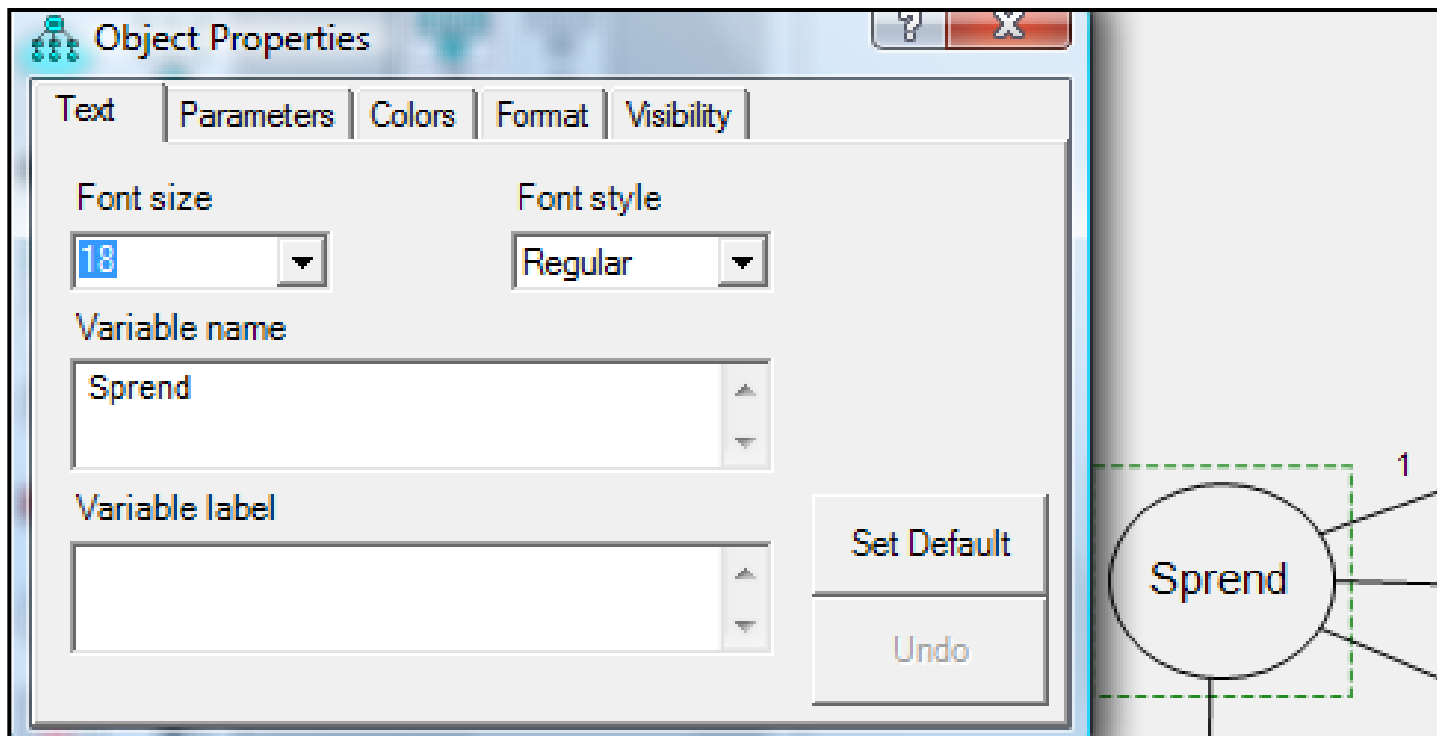
Diagramos braižymas

Ant pirmosios elipsės spragtelime dešiniuoju klavišu ir renkamės viršutinę ikonėlę (AMOS 5 reikėtų dukart spragtelėti)



Diagramos braižymas

Pasirinkę *Text* įrašome vardą *Sprend*.

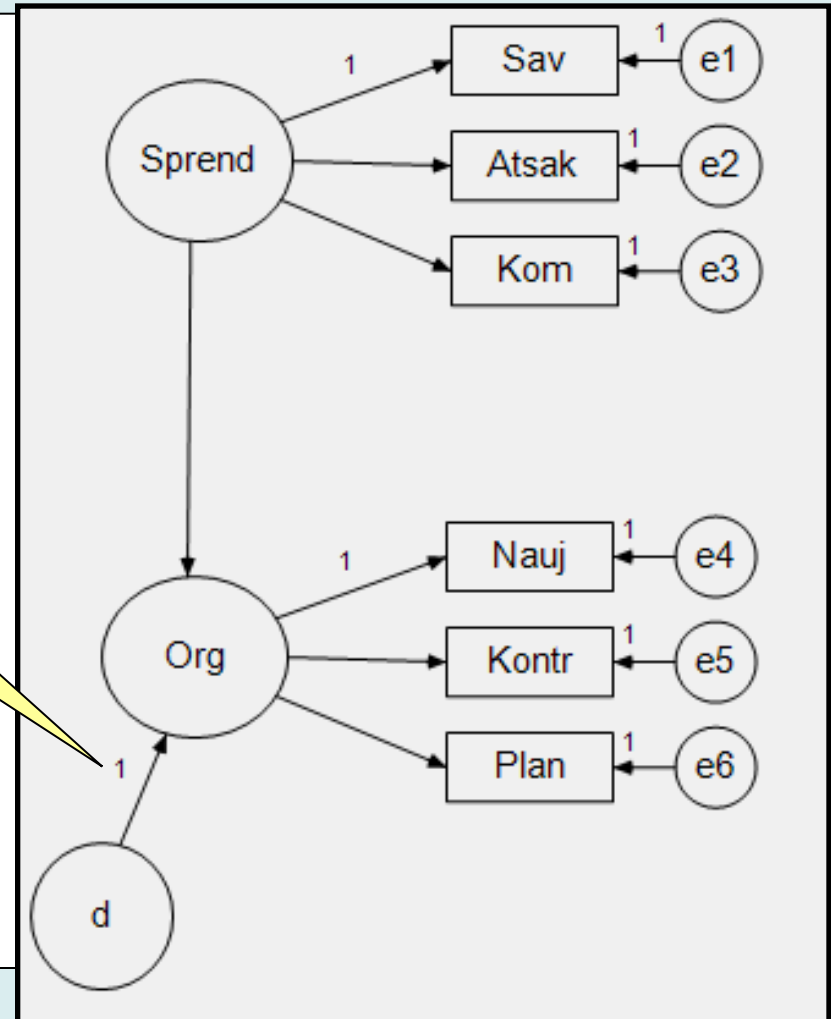


Diagramos braižymas

Neužvėrę langelio,
įvardijame kitą faktorių
Org ir paklaidas
 $d, e1, \dots, e6$.

Svoris 1

Uždveriname
vardijimo langelį.



Diagramos braižymas

Parodėme tik vieną iš galimų diagramos braižymo būdų. Yra ir kitų galimybių.

Pavyzdžiui, galima automatiškai suteikti vardus liekamosioms paklaidoms arba aprašyti pavadinimą; pagrindines charakteristikas (chi kvadrato p reikšmę ir kt.) užrašyti virš brėžinio ir pan.

Parinkčių nustatymas

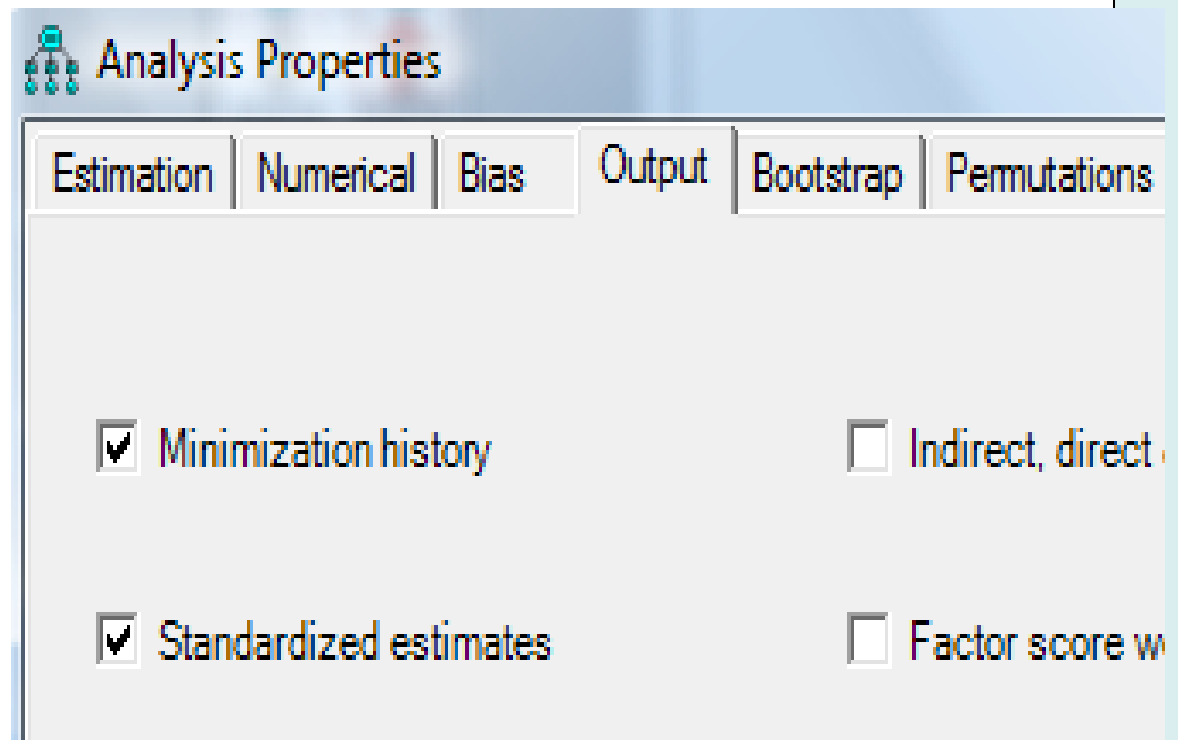
Pasirenkame



. Pasirenkame *Output*.

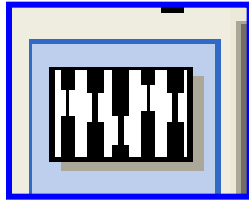
Pažymime:

Standardized Estimates.



Programos vykdymas

Spragtelime



Failą išsaugome (jei sunku sugalvoti failo vardą, tai pavadinkite jį A1).

Tada spragtelime



(*view text*).

Rezultatų analizė

The screenshot shows a software interface with a left-hand navigation pane and a main content area. The navigation pane includes the following items: Parameter summary, Notes for Model (highlighted with a blue box), Estimates, Minimization History, Model Fit, and Execution Time. The main content area displays the following text:

Number of distinct sam
Number of distinct parameters to
Degrees of freed

Result (Default model)

Minimum was achieved
Chi-square = 5.399
Degrees of freedom = 8
Probability level = .714

A yellow callout bubble points to the underlined value ".714" in the "Probability level" line. Inside the bubble, the text reads: "P reikšmė 0.714".

Chi kvadrato kriterijus rodo gerą bendrąjį modelio tinkamumą duomenims ($p > 0,05$).

Rezultatų analizė

Nereikia suabsoliutinti chi kvadrato kriterijaus svarbos. Jeigu duomenų bus labai daug (pvz., 2000), tai statistiškai reikšminga chi kvadrato statistika dar nieko nereiškia. Reikia pasižiūrėti, ar pati statistikos reikšmė nėra didelė (nagrinėjamame pavyzdyje $\chi^2 = 5,399$). Nagrinėjant dideles imtis, svarbiau kiti geri tinkamumo duomenims rodikliai: *RMSEA*, *GFI* ir pan.

Bendrieji rodikliai

GFI ir
RMSEA rodo
gerą bendrąjį
modelio
tinkamumą
duomenims.

Model	RMR	GFI	AGFI	PGFI
Default model	.098	.988	.969	.376
Saturated model	.000	1.000		
Independence model	.932	.612	.457	.437

Model	RMSEA
Default model	.000
Independence model	.304

Kelių koeficientai

Kelių koeficientai

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
Org	<---	Sprend	.800	.249	3.221	.001	
Sav	<---	Sprend	1.000				
Atsak	<---	Sprend	.516	.181	2.843	.004	
Kom	<---	Sprend	.566	.283	2.000	.045	
Nauj	<---	Org	1.000				
Kontr	<---	Org	2.572	.313	8.230	***	
Plan	<---	Org	2.095	.276	7.576	***	

Visi kelių koeficientai statistiškai reikšmingi.

Rezultatų analizė

Statistiškai reikšmingas kelio koeficientas rodo, kad modelyje kelias tinka. Jeigu kelio koeficientas statistiškai nereikšmingas, tai modelį reikėtų tobulinti.

Patys kelių koeficientai dar neleidžia įvertinti, kuris kelias (spėjama priklausomybė) svarbesnis. Tą galima padaryti, peržiūrėjus standartizuotuosius kelių koeficientus.

Standartizuotieji koeficientai

Standardized Regression Weights: (Group number 1 - Default model)

		Estimate
Org	<--- Sprend	.829
Sav	<--- Sprend	.738
Atsak	<--- Sprend	.298
Kom	<--- Sprend	.201
Nauj	<--- Org	.692
Kontr	<--- Org	.866
Plan	<--- Org	.721

Stipriausia priklausomybė tarp *Org* ir *Kontr*.

Dispersijų įverčiai

Įverčiai

Statistinis
reikšmingumas

Variance (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
Sprend	.615	.211	2.914	.004	
d	.179	.122	1.471	.141	
e1	.515	.185	2.780	.005	
e2	1.677	.203	8.254	***	
e3	4.677	.554	8.437	***	
e4	.622	.089	6.997	***	
e5	1.268	.338	3.750	***	
e6	2.324	.348	6.674	***	

Dispersijų įverčiai

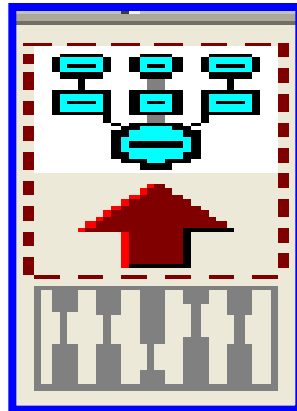
Didelės dispersijos rodo, kad modelis nėra labai geras (lika daug nepaaiškintos informacijos).

Esant statistiškai nereikšmingai dispersijai, reikėtų patikrinti, ar modelyje reikalinga atitinkama liekamoji paklaida. Vis dėlto skubėti atsisakyti liekamosios paklaidos nereikėtų. Liekamosios paklaidos nebuvimas rodo, kad kintamojo elgesį 100 % paaiškina jo regresoriai, o taip būna labai retai.

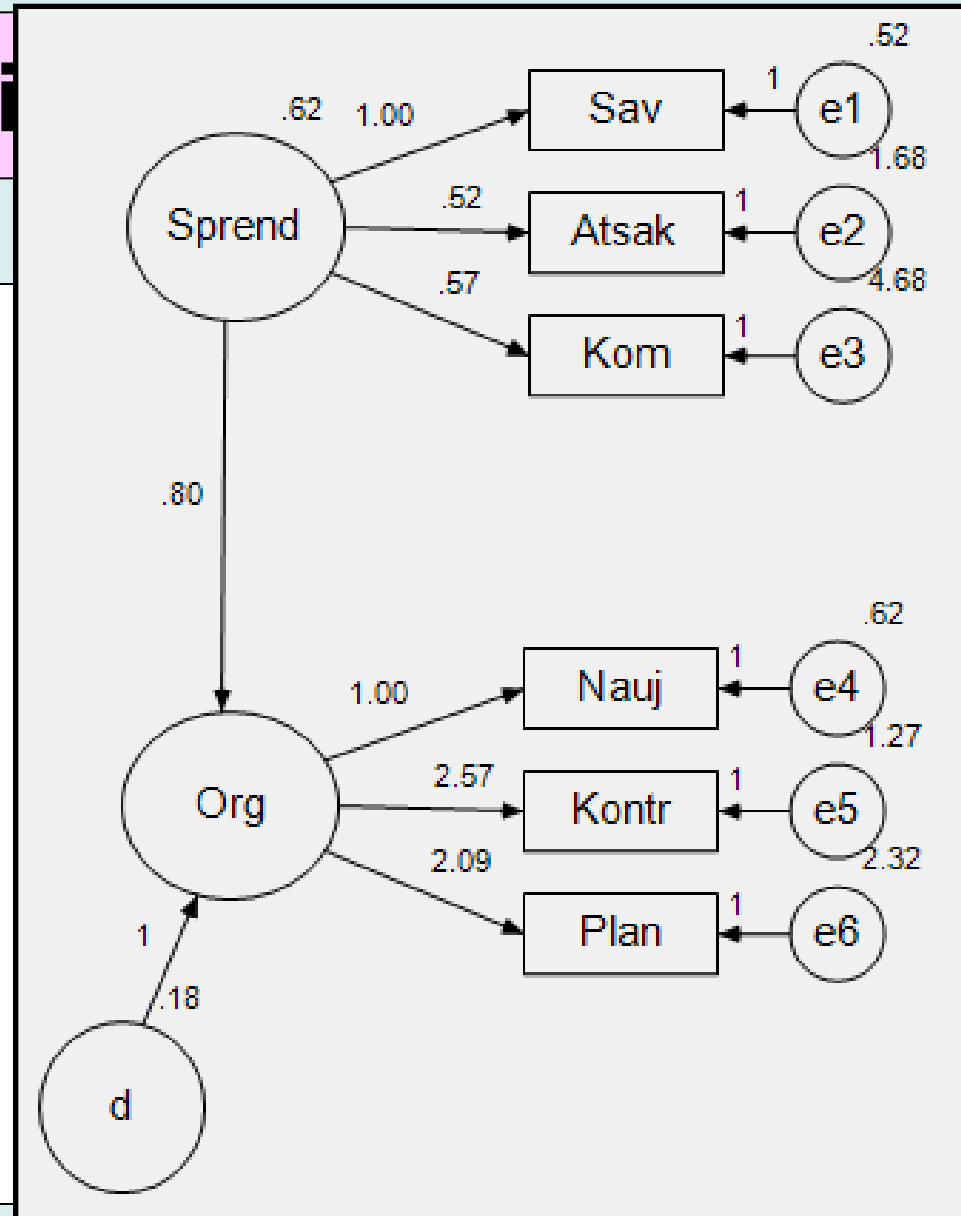
Mes liekamąją paklaidą d modelyje paliekame.

Rezultatų analizė

Norėdami modelyje pamatyti parametrų įverčius ir dispersijas, spragtelime

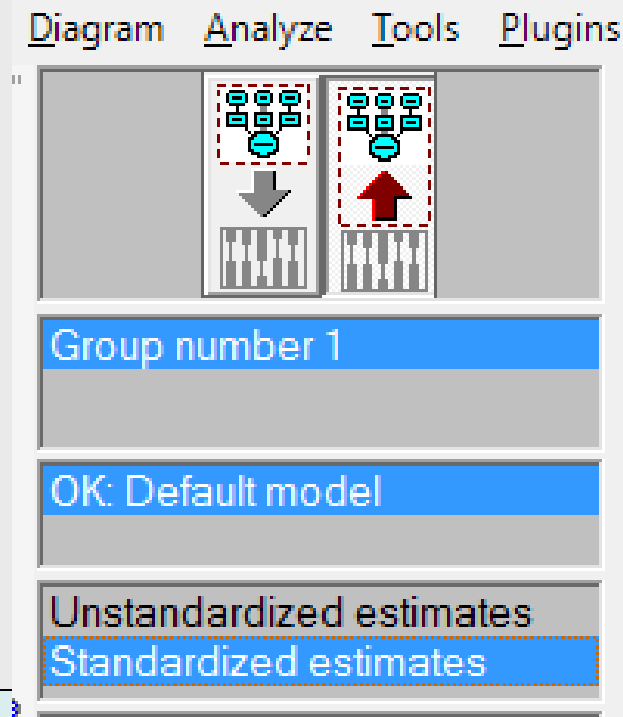


Parametrų įverčiai

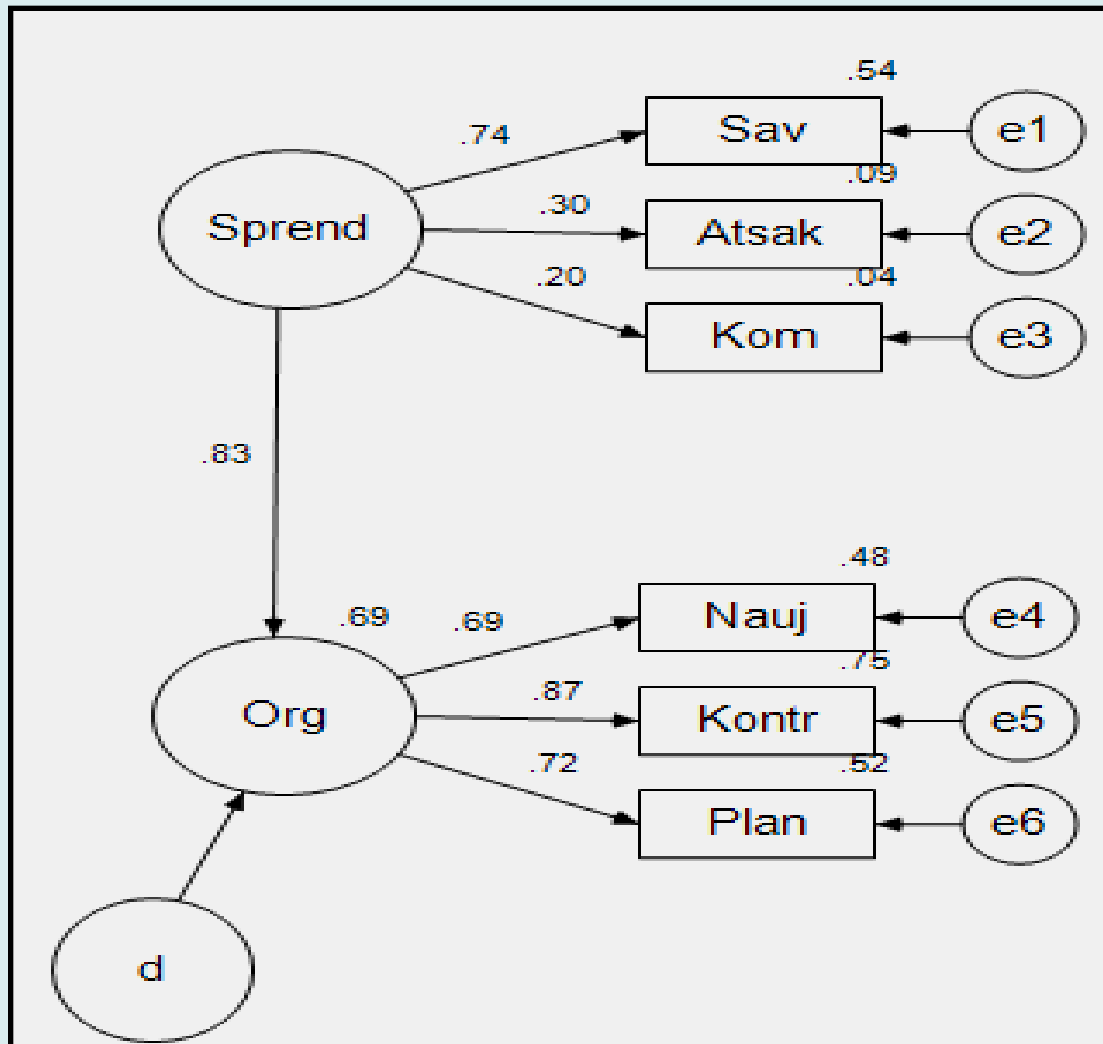


Standartizuotieji kelių koeficientai

Spragtelėjus *Standardized estimates*, gaunami standartizuotieji kelių koeficientai.



Standartizuotieji kelių koeficientai



Galutinė išvada

Turimi duomenys neprieštarauja
spėjamamoms priklausomybėms.

Žinoma, tai dar nereiškia, kad negalima
sudaryti kitų modelių.

Pastabos

Su AMOS programa galima atlikti ir automatinę kelių paiešką.

Be to, galima fiksuoti dalį parametrų.

Galima fiksuoti ne pačias parametrų reikšmes, o jų santykį. Pavyzdžiui, numatyti, kad kelių *Sprend* \rightarrow *Atsak* ir *Sprend* \rightarrow *Kom* koeficientai lygūs.