

ZÁKLADY EKONOMETRIE (KOMBINOVANÁ FORMA) – SKUPINOVÝ ÚKOL 2

Pro účely zpracování úkolu si vytvořte pracovní skupinu o max. 5 lidech. Úkoly zpracujte a odevzdejte do stanoveného termínu. Veškeré podklady (např. relace gretlu) a komentáře, které mohou být i součástí relace gretlu, odevzdejte do příslušné odevzdárny.

Termín odevzdání: 7. 12. 2015 (včetně)

Zadání úkolu

1. (Hill et al. (2011): př. 8.13, 8.14) Uvažujme následující nákladovou funkci Firmy 1 z odvětví výroby oblečení, kde C označuje náklady a Q výstup. Předpokládejme, že $\text{var}(\epsilon_{1t}) = \sigma^2 Q_{1t}$. Používáme index t z důvodu, že máme pozorování v podobě časových řad (celkem 28 pozorování). Data jsou obsahem souboru `cloth.gdt` (pokud instalace gretlu datové soubory z nového vydání Hill et al. (2011) neobsahuje, je datový soubor součástí podkladových materiálů ke cvičení).

$$C_{1t} = \beta_1 + \beta_2 Q_{1t} + \beta_3 Q_{1t}^2 + \beta_4 Q_{1t}^3 + \epsilon_{1t}.$$

- (a) Nalezněte odhady parametrů $\beta_1, \beta_2, \beta_3$ a β_4 metodou zobecněných nejmenších čtverců.
- (b) Testujte hypotézu $\beta_1 = \beta_4 = 0$. Co můžete říct o povaze funkce průměrných nákladů, pokud je tato hypotéza pravdivá?
- (c) Za jakých předpokladů (kladených na náhodnou složku) by bylo vhodnější odhadnout funkci průměrných nákladů než funkci celkových nákladů?
Uvažujme nyní, že nákladová funkce i druhé firmy (Firma 2) v tomto odvětví je kubická. Máme tedy následující nákladové funkce:

$$\text{Firma 1 : } C_{1t} = \beta_1 + \beta_2 Q_{1t} + \beta_3 Q_{1t}^2 + \beta_4 Q_{1t}^3 + \epsilon_{1t},$$

$$\text{Firma 2 : } C_{2t} = \delta_1 + \delta_2 Q_{2t} + \delta_3 Q_{2t}^2 + \delta_4 Q_{2t}^3 + \epsilon_{2t},$$

kde $E(\epsilon_{1t}) = E(\epsilon_{2t}) = 0$, $\text{var}(\epsilon_{1t}) = \sigma_1^2$ a $\text{var}(\epsilon_{2t}) = \sigma_2^2$. Předpokládáme, že ϵ_{1t} a ϵ_{2t} jsou vzájemně nezávislé.

- (d) Odhadněte každou z funkcí metodou nejmenších čtverců. Komentujte výsledky. Jsou znaménka odhadnutých parametrů v souladu s očekáváním?
- (e) Na hladině významnosti 10 % testujte hypotézu $H_0 : \sigma_1^2 = \sigma_2^2$ oproti alternativě $H_1 : \sigma_1^2 \neq \sigma_2^2$ (použijte standardní F -test shody rozptylů).
- (f) Odhadněte obě rovnice společně za předpokladu, že $\beta_1 = \delta_1, \beta_2 = \delta_2, \beta_3 = \delta_3$ a $\beta_4 = \delta_4$. Komentujte dosažené výsledky.

[Nápověda: pokud máme ve vzorku dvě proměnné x_1 a x_2 , potom pomocí funkce `stack` v rámci příkazu `Přidat > Definovat novou proměnnou...` a zápisem `x=stack(x1, x2)` nebo v konzoli gretlu příkazem `genr x=stack(x1, x2)` vytvoříme novou proměnnou x , která bude obsahovat obě předchozí proměnné poskládané za sebou. Délka datového vzorku se automaticky zvětší na dvojnásobek. Proměnná se nemusí v seznamu proměnných zobrazit hned, nicméně je vytvořena a k dispozici pro odhad. Ke grafické aktualizaci dojde, když např. vytvoříme novou proměnnou „standardním“ způsobem, bez použití funkce `stack`, např. jako součet dvou proměnných nebo nahrazením existující proměnné.]

- (g) Testujte hypotézu

$$H_1 : \beta_1 = \delta_1, \beta_2 = \delta_2, \beta_3 = \delta_3, \beta_4 = \delta_4$$

a komentujte výsledek testu.

2. (Ramantahan (2002): př. 9.22) Datový soubor `data9-8.gdt`, či položka `data9-8` na záložce Ramanathan v gretlu obsahuje roční data z oblasti letecké dopravy ve Spojených státech, a to konkrétně data následujících proměnných:

- pop , počet obyvatel v milionech,
- rpm , příjem z osobomil (domácích) v miliardách dolarů,
- nop , počet leteckých společností (provozovatelů letecké přepravy),
- $oprev$, provozní příjem z přepravených pasažérů v miliardách dolarů,
- gnp , hrubý národní produkt Spojených států v miliardách dolarů,
- $accid$, počet nehod amerických dopravních letadel,
- $fatal$, počet zemřelých při leteckých neštěstích,
- $regu$, umělá proměnná pro deregulaci v letecké přepravě (0 je pro roky 1945 až 1978 a 1 je pro období po roku 1979).

(a) Odhadněte následující model metodou nejmenších čtverců a interpretujte výsledky:

$$\ln\left(\frac{rpm}{pop}\right) = \beta_1 + \beta_2 regu + \beta_3 \ln\left(\frac{oprev}{rpm}\right) + \beta_4 \ln\left(\frac{gdp}{pop}\right) + \beta_5 \ln\left(\frac{accid}{rpm}\right) + \beta_6 fatal + u,$$

kde $\frac{rpm}{pop}$ je příjem s osobomil na obyvatele, $\frac{gnp}{pop}$ je důchod na hlavu, $\frac{oprev}{rpm}$ vyjadřuje průměrný příjem z jené míle a $\frac{accid}{rpm}$ je míra nehodovosti na osobomíli. Jedná se o log-log model s výjimkou proměnných $regu$ a $fatal$ (nabývají i nulových hodnot).

- (b) Testuje autokorelaci reziduí, a to autokorelaci prvního řádu. Co můžete říct na základě získaných výsledků o vlastnostech odhadů, kvalitě vyrovnání a testových statistikách (výsledcích testů hypotéz)?
- (c) Odhadněte model Cochrane-Orcuttovu procedurou případně využijte HAC estimátor. Porovnejte výsledky s výsledky předchozími.
- (d) Předpokládejme, že chceme testovat to, jestli úrovně konstanta a koeficienty mezních vlivů (z nichž některé vyjadřují elasticity) závisí na proměnné $regu$. Jinými slovy, předpokládejme, že zde může docházet ke strukturální změně v důsledku deregulace odvětví letecké přepravy. Formulujte tedy (obecný) model s dodáním dodatečných interakčních členů a naznačte, co je myšleno nulovou hypotézou, která říká, že k žádné strukturální změně nedochází. Odhadněte odpovídající modely pro testování této hypotézy a test proveďte.
- (e) Využijte Cochrane-Orcuttovu proceduru k získání vydatnějších odhadů obecného modelu a s využitím vhodných testů jej upravte do podoby modelu, ve kterém budou všechny koeficienty statisticky významné (a tedy modelu s vydatnějšími odhady).
- (f) Otestujte původní (obecný) model na přítomnost autokorelace třetího řádu a v případě nutnosti využijte zobecněnou C-O proceduru zohledňující tuto specifikaci náhodných složek. Jak by tato specifikace rovnicově vypadala?
3. (Hill (2008), př. 10.9) Uvažujme model nabídky kuřat pěstovaných pro maso, které americké ministerstvo zemědělství (a nejen to) nazývá brojler. Empirická data jsou obsahem souboru `newbroiler.gdt` (dostupný skrze záložku `POE`), který byl poskytnut Epplem a McCallumem (2006), autory článku „Simultaneous equation econometrics: The missing example,“ z časopisu *Economic Inquiry*, 44(2), str. 374-384. Data jsou roční z období 1950–2001, nicméně v odhadu využijte data jen pro období 1960–1999. Rovnice nabídky je dána jako

$$\ln(QPROD_t) = \beta_1 + \beta_2 \ln(P_t) + \beta_3 \ln(PF_t) + \beta_4 TIME_t + \beta_5 \ln(QPROD_{t-1}) + \epsilon_t^s,$$

kde $QPROD$ = agregátní produkce mladých kuřat, P = reálný cenový index čerstvých kuřat, PF = reálný cenový index krmiva pro brojler, $TIME = 1, \dots, 52$. Tato rovnice nabídky je dynamická, se členem zpožděné produkce na pravé straně. Tuto proměnnou budeme chápat jako známou a predeterminovanou v čase t . Časový index je zahrnut pro zachycení technologického pokroku v produkci. Možné instrumentální proměnné jsou

$\ln(Y_t)$, kde Y je důchod na hlavu, $\ln(PB_t)$, kde PB je reálná cena hovězího, dále $POPGRO$, což je procento tempo růstu populace mezi lety $t - 1$ a t , $\ln(P_{t-1})$ jako logaritmus zpožděné reálné ceny kuřat, a konečně $\ln(EXPTS)$, což je logaritmus exportů kuřat.

- (a) Odhadněte rovnici nabídky metodou nejmenších čtverců. Diskutujte získané výsledky. Jsou znaménka odhadů a významnost jednotlivých koeficientů v souladu s vaším očekáváním?
- (b) Odhadněte rovnici nabídky s využitím estimátoru instrumentálních proměnných. Pracujeme s modelem nabídky a poptávky a je tak velmi pravděpodobná endogenita ceny brojlerů (cenu krmiva chápeme jako exogenní veličinu). K odhadu využijte všechny dostupné instrumenty. Porovnejte získané výsledky s těmi získanými prostým OLS odhadem.
- (c) Testujte endogenitu $\ln(P_t)$. Test proveďte jakožto Hausmannův test. Jaký je závěr vašeho testu?
- (d) Otestujte sílu instrumentů s využitím standardní F -statistiky pomocné regrese endogenní proměnné na vaše všechny instrumenty (první fáze TSLS). Jaký závěr můžete učinit? V rámci testování můžeme od nancí výsledného rozdělení této statistiky abstrahovat (díky nestandardnímu rozdělení gretl nabízí speciální tabelované kritické hodnoty, případně pro více endogenních proměnných Cragg-Donaldovu statistiku). Hrubým pravidlem je to, aby příslušná F -statistika byla vyšší než 10, popř. jednotlivé t -statistiky byly vyšší než 3.3. Přestože se dělá pomocná regrese na všechny instrumenty (kterými jsou i exogenní vysvětlující veličiny původní regrese), v rámci vyhodnocení F -testu se díváme na společnou (či u t -testu individuální) statistickou významnost jen instrumentů bez těchto exogenních veličin původní regrese.
- (e) Má volba instrumentů svůj logický základ? Pokud ano, proč? Otestujte validitu vašich instrumentů (samostatně, nebo s využitím výsledků Sarganova testu).
- (f) Zvažte sami, jestli kromě ceny kuřat nechápat i cenu krmiva (a samozřejmě i její logaritmus) jako potenciálně endogenní veličinu a případně podle toho upravte předchozí odhady.