

Sensorveiledning SØK 3001 V16

Oppgave 1.

a) Formuler markedsmodellen for trøfler basert på disse opplysningene og forklar hvorfor OLS-estimatoren for etterspørsels og tilbudsfunksjonen for trøfler vil være skjev.

Underliggende markedsmodell, forutsatt markedsklarering, etterspørsel lik tilbud

$$(1) q_i = \alpha_0 + \alpha_1 P_i + \alpha_2 PS_i + \alpha_3 DI_i + u_{di} \quad \text{Etterspørselsfunksjon}$$

$$(2) q_i = \beta_0 + \beta_1 P_i + \beta_2 PF_i + u_{si} \quad \text{Tilbudsfunksjon}$$

Poenget er å innse at prisen P og restleddene i strukturrelasjonene (1) og (2) er korrelerte. Kan resonnerer på markedsmodellen grafisk og overbevise leseren om at restleddet i etterspørselsfunksjonen og tilbudsfunksjonen er korrelert med restleddet rent intuitivt. Effektive kandidater vil eventuelt også formulere redusert form-ligninger for likevektspris og likevektskvantum som de får bruk for senere og bruke prisligninga til å vise at prisen P og restleddene i etterspørsels og tilbudsfunksjonen vil være korrelerte og at det gir skjevhet i OLS-estimatoren.

Redusert form-ligninger: Løser (1) og (2) mhp de endogen q og P, gir

$$(3) q_i = \pi_{0q} + \pi_{1q} PS_i + \pi_{2q} DI_i + \pi_{3q} PF_i + e_{qi}$$

$$(4) p_i = \pi_{0p} + \pi_{1p} PS_i + \pi_{2p} DI_i + \pi_{3p} PF_i + e_{pi}$$

$$\text{Her er } e_{pi} = \frac{u_{di} - u_{si}}{\alpha_1 - \beta_1} \quad \text{og} \quad e_{qi} = \frac{\alpha_1 u_{di} - \beta_1 u_{si}}{\alpha_1 - \beta_1}$$

Fra (4) ser vi at u_{di} og P er positivt korrelert, mens u_{si} og P er negativt korrelert. Med andre ord er det brudd på en sentral betingelse for at OLS-estimatoren skal være konsistent.

b) Gi en tolking av de estimerte koeffisientene i Tabell 1.

Estimerte koeffisienter i tabell 1 er koeffisienter i redusert form ligninga for henholdsvis likevektsprisen P og likevektskvantum q som lineære funksjoner av de eksogene variablene i modellen.

c) Forklar hvordan du ville benyttet informasjonen i Tabell 1 til å estimere etterspørsels og tilbudsfunksjonen for trøfler?

Rett fram to-steps 2SLS (Woolridge kap om IV)

Prisligninga (reduisert form) i tabell 1 representerer 1. steget. Erstatt predikert verdi for prisen fra redusert-form ligninga for likevektsprisen i 2. steget og estimer den da framkomne ligning med OLS.

d) Forklar hvordan du kan benytte informasjonen i Tabell 1 til å undersøke om etterspørsels og tilbudsfunksjonen er identifisert i empirisk forstand?

I følge modellen vil timeprisen på leie av letegriser PF bidra til å identifisere etterspørselsfunksjonen. Etterspørselsfunksjonen er dermed eksakt identifisert.

Etterspørselsfunksjonen er identifisert i empirisk forstand dersom PF bidrar signifikant til å forklare variasjon i likevektsprisen (gyldig instrument). Det svarer til å teste nullhypotesen om at koeffisienten foran PF i prisligninga i tabell 1 er lik null. Her ser vi at koeffisienten er 1,35 med standardavvik 0.298, tilhørende t-verdi innebærer at nullhypotesen forkastes med klar margin hvilket indikerer at etterspørselsfunksjonen er identifisert i empirisk forstand.

Vi har på samme måte to variable, PS og DI som ifølge modellen bidrar til å identifisere tilbudsfunksjonen. Tilbudsfunksjonen er dermed overidentifisert. Tilbudsfunksjonen kan sies å være identifisert i empirisk forstand dersom PS og DI bidrar signifikant til å forklare variasjon i likevektsprisen (gyldige instrumenter). Koeffisientene foran PS og DI i prisligninga i tabell 1 er h.h.v. 1.708 og 7.603 med tilhørende standardavvik 0.351 og 1.724. Tilhørende t-verdier innebærer at vi klart kan forkaste separate nullhypoteser om at hver av dem er null. Vi kunne også ønske å teste den simultane hypotesen om at koeffisientene foran PS og DI begge er lik null med F-test, men det har vi ikke informasjon til å kunne gjøre med opplysningene som er gitt i tabell 1.

e) Tabell 2 viser estimerte tilbuds- og etterspørselsfunksjoner for trøfler. Tolk de estimerte koeffisientene. En kommentator påstår at uekte trøfler er et substitutt til ekte trøfler. Hvordan vil du gå fram for å teste denne påstanden? Hvordan vil du tolke den estimerte effekten av letekostnadene på tilbudet av trøfler?

Estimatene foran P viser at når prisen på trøfler øker med 1 dollar så reduseres etterspørselen med 0.37 unser. Effekten er statistisk signifikant (tilhørende standardavvik er 0.1648. t-verdien er 2.27 og over kritisk verdi). Effekten av økt inntekt, DI er positiv med estimat på

5.01. Standardavvik 2.28 innebærer at inntektseffekten er signifikant ulik null (t-test), og estimatet innebærer at en økning i månedsinntekta med 1000 dollar gir en økning i etterspørselen etter trøfler med 5 unser.

Hvis trøfler og uekte trøfler er substitutter vil koeffisienten foran PS i etterspørselsfunksjonen være positiv, altså at etterspørselen etter ekte trøfler øker dersom prisen på substituttet (uekte trøfler) øker. Kan teste påstanden ved å teste nullhypotesen $\alpha_2 = 0$ i (1) mot alternativhypotesen om at $\alpha_2 > 0$ ved ensidig t-test. Koeffisienten er estimert til 1.2960 og tilhørende standardavvik er 0.3522. t-observator er 3.7 og klart over kritisk grense i t-fordelinga. Resultatene tyder derfor på at kommentatorens påstand er riktig. En økning i prisen på uekte trøfler på 1 dollar øker etterspørselen etter ekte trøfler med 1.3 unser. Effekten av Letekostnadene PF på tilbudet er estimert til -1.00 og tilhørende standardavvik er 0.083. Estimatet (som er klart signifikant ulikt null basert på tosidig t-test) innebærer at en økning i letekostnadene med 1 dollar per time reduserer tilbudet av trøfler med 1 unse. Et negativt estimat er som ventet og i tråd med grunnleggende tilbudsteori. PF er en kostnadsfaktor i produksjonen av trøfler og økt PF gir et negativt horisontalt skift i tilbudskurven.

f) En student foreslår at det bør gjennomføres en test av overidentifiserende restriksjoner for å undersøke troverdigheten av estimeringsresultatene i Tabell 2. Kommenter dette forslaget.

Bør kommentere at test av overidentifiserende restriksjoner bare er mulig på relasjoner som er overidentifiserte, i den forstand at det er to potensielle instrumenter tilgjengelig. Her gjelder det bare tilbudsfunksjonen. Etterspørselsfunksjonen er som beskrevet ovenfor eksakt identifisert (1 tilgjengelig instrument). Prosedyren for testen er kort fortalt å estimere en hjelpregresjon med OLS, der 2SLS-residualene er avhengig variabel og alle de eksogene variablene i strukturligninga (her: PF) samt instrumentene for P i tilbudsligninga (PS og DI), inngår som forklaringsvariable. Under nullhypotesen om at alle instrumentene er ukorrelerte med restleddet u_{si} så vil observatoren $N \times R^2$ der N er antall observasjoner og R^2 er multiple determinasjonskoeffisient i hjelpregresjonen følge en kjikvadrat-fordeling med 1 frihetsgrad. Hvis testobservatoren overstiger kritisk verdi i kjikvadratfordelinga med 1 frihetsgrad så forkastes hypotesen. Se lærebok kap 15.5 b for diskusjon av prosedyren og intuisjonen bak testen.

Oppgave 2.

Relevant å spesifisere modellen som er brukt i de ulike kolonnene.

$$(1) b_t = \beta_0 + \beta_1 p_t + \beta_2 y_t + u_t$$

$$(2) b_t = \beta_0 + \alpha b_{t-1} + \beta_{10} p_t + \beta_{11} p_{t-1} + \beta_{20} y_t + \beta_{21} y_{t-1} + u_t$$

$$(3) \Delta b_t = \beta_0 + (\alpha - 1) b_{t-1} + \beta_{10} \Delta p_t + (\beta_{10} + \beta_{11}) p_{t-1} + \beta_{20} \Delta y_t + (\beta_{20} + \beta_{21}) y_{t-1} + u_t$$

a) Hva er tolkingen av koeffisientene i kolonne (1)?

(1) er en statisk modell der det forutsettes at den fulle effekten av pris og inntekt slår inn i inneværende periode. Resultatene innebærer en priselastisitet på -0.415 og en inntektselastisitet på 0.467. 1% økning i realprisen på gir dermed drøyt 0.4 prosent reduksjon i salget, mens 1% økning i realinntekta gir 0.47 prosent økning i bensinsalget.

b) Bruk resultatene i kolonne (2) til å beregne kortsikts og langsikts pris og inntektselastisitet i bensinsalget.

Modellen i kolonne (2) og (3) er dynamiske representasjoner av modellen for bensinsalg. Kandidatene bør her innse at modellen i kolonne (3) er en ren transformasjon(reparametrisering) av modellen i kolonne (2)-altså samme underliggende modell. I (2) og (3) tillates bensinsalget å reagere på endringer i realpris og realinntekt med tidsetterslep, noe som kan være en mer realistisk forutsetning enn (1). Kommer tilbake til dette i spørsmål c). Korttids priselastisitet blir -0.305, mens korttids inntektselastisitet blir 0.194.

Langtidspriselastisitet finnes ved å anta en hypotetisk langtidslikevekt der $b_t = b_{t-1} = b$,

$$p_t = p_{t-1} = p \text{ og } y_t = y_{t-1} = y$$

Da blir langtidspriselastisitet lik $(\beta_{10} + \beta_{11}) / (1 - \alpha) = -0.084 / 0.127 = -0.661$, mens langtidsinntektselastisitet blir $(\beta_{20} + \beta_{21}) / (1 - \alpha) = -0.022 / 0.127 = -0.173$

c) En student foreslår at den estimerte relasjonen i kolonne (2) bør foretrekkes framfor relasjonene i kolonne (1) og (3) basert på den multiple determinasjonskoeffisienten R^2 . Kommenter dette forslaget.

Kandidatene bør innse at det bare er meningsfylt å sammenligne R^2 for modeller med samme avhengige variabel. Her kan (1) og (2) sammenlignes, og ikke uventet er den dynamiske modellen (2) å foretrekke, selv om det her kan kommenteres at en sammenligning basert på justert R^2 ville vært å foretrekke siden antall variable som inngår i (2) er høyere enn i (1). Mer presis sammenligning fåes ved å bruke F-test siden (1) er et spesialtilfelle av (2). F-test basert på R^2 blir

$$F^{obs} = \frac{(R_{(2)}^2 - R_{(1)}^2) / q}{(1 - R_{(2)}^2) / (n - k - 1)} = \frac{(0.98 - 0.47) / 3}{(1 - 0.98) / 29 - 6} = 195.5$$

Siden F-observatoren er klart over kritisk grense i F-fordelinga med 3 og 23 frihetsgrader forkastes hypotesen om at (1) er en gyldig forenkling av (2). (3) er bare en ren transformasjon/reparametrisering av (2) og det er meningsløst å skille mellom disse ekvivalente modellene.

d) Bruk resultatene i tabell 3 til å teste en hypotese om at den langsiktige inntektselastisiteten er lik null.

Det følger av definisjonen av langtidselastisiteter gitt ovenfor og formuleringen av modellen som (3) at en test av hypotesen om langtidsinntektselastisitet lik null er ekvivalent med å teste med t-test om koeffisienten foran y_{t-1} i kolonne (3) er ulik null, dvs. $(\beta_{20} + \beta_{21})=0$. t-verdien her er -0.82. Klart lavere enn kritisk verdig i t-fordelinga og vi kan ikke forkaste hypotesen om at langtids inntektselastisitet er lik null.

e) Basert på residualene fra regresjonen i kolonne (2) er det beregnet en Breush-Godfrey LM-testobservator for å teste for første og andre ordens seriekorrelasjon i restleddet. Verdien på testobservatoren er 0.681. Forklar hvordan denne observatoren er beregnet og bruk informasjonen til å teste hypotesen om at restleddet ikke er seriekorrelert.

Seriekorrelasjon forklares kort og det bør formuleres en modell for restleddsprosessen. Anta restleddet u følger en AR(2) prosess, $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t$

Test om fravær av 1. og 2 ordens seriekorrelasjon innebærer en teste hypotesen $\rho_1=0$ og $\rho_2=0$

Testen implementeres ved å kjøre hjelperegresjon mellom residualen fra (2) mot residualen lagga 1 og 2 perioder og forklaringsvariablene i modellen. Kan gjøres som en LM-test (Breush&Godfrey). Kan vises at $(N-2) \times R^2$ fra hjelperegresjonen (N er antall observasjoner) følger en kjikvadratfordeling med 2 frihetsgrader under nullhypotesen. Siden 0.681 er klart lavere enn kritisk verdi i kjikvadratfordelinga

med 2 frihetsgrader kan vi ikke forkaste hypotesen om fravær av 1. og 2. ordens seriekorrelasjon i modellen i kolonne (2).

f) En kommentator foreslår at det foretas en utvidelse av modellen estimert i kolonne (3) som kan brukes til å teste en hypotese om at bensinsalgsrelasjonen er homogen av grad null i nominelle størrelser både på kort og lang sikt. Formuler en slik utvidet modell og forklar hvordan du ville gå fram for å teste hypotesen.

Utvidelse, testing av homogenitet av grad 0: Kan bruke en utvidet versjon av den reparametriserte dynamiske-modellen i kolonne (3) til å teste homogenitet av grad 0 på kort og lang sikt

Hittil: I de estimerte relasjoner (1)-(3) er det pålagt restriksjon om homogenitet av grad 0 i nominelle priser og inntekt siden det bare er realpriser og realinntekter som inngår.

Kan formulere relasjon uten denne restriksjonen ved å utvide relasjonen (3) som i utgangspunktet er formulert med realpris og realinntekt på høyresida med ledd i logaritmen til konsumprisindeksen, kpi_t . Her vil 1.differens i logaritmen til konsumprisindeksen, Δkpi_t representere avvik fra homogenitet av grad 0 på kort sikt, mens lagga logaritme til konsumprisindeksen kpi_{t-1} , representere avvik fra homogenitet av grad 0 på lang sikt. Relasjon (4) viser en slik utvidet variant.

$$(4) \Delta b_t = \beta_0 + (\alpha - 1)b_{t-1} + \beta_{10}\Delta \ln p_t + (\beta_{10} + \beta_{11})p_{t-1} + \beta_{20}\Delta y_t + (\beta_{20} + \beta_{21})y_{t-1} + \gamma_0\Delta kpi_t + (\gamma_0 + \gamma_1)kpi_{t-1} + u_t$$

Restriksjon om homogenitet av grad 0 på både kort og lang sikt innebærer:

$$\gamma_0 = 0 \text{ og } (\gamma_0 + \gamma_1) = 0$$

Kan testes ved F-test. Forkast H_0 dersom beregna F-verdi overstiger kritisk verdi i F-fordelinga med relevante antall frihetsgrader i teller og nevner.

Restriksjon om homogenitet av grad 0 på lang sikt, men ikke på kort sikt:

$$(\gamma_0 + \gamma_1) = 0$$

Kan testes med t-test på koeffisienten foran kpi_{t-1} i modell (4)

Oppgave 3.

En kritisk kommentator kommer med følgende påstand: «Økonometrikerne estimerer sammenhenger mellom variable som inneholder betydelige målefeil og derfor kan vi ikke stole på resultatene fra økonometriske studier». Drøft denne påstanden med grunnlag i økonometrisk teori.

Standard diskusjon om målefeil og at problemet med målefeil er avhengig av hvilke(n) variabel som inneholder målefeil samt korrelasjonen mellom målefeilen og den observerbare variabelen. Skille mellom tre typer målefeil og diskutere problemet basert på en enkel økonometrisk ligning. Observert Y vs teoretisk Y^* , Observert X versus teoretisk X^*

1) målefeil i avhengig variabel, y (Woolridge, kap 9-4.a), OLS fortsatt konsistent, høyere varians, mindre presis OLS-estimator

2) Målefeil i forklaringsvariabel, der målefeilen er ukorrelert med observert variabel, X (Woolridge kap 9-4.b). Samme konklusjon som a), OLS fortsatt konsistent

3) Klassisk målefeilmodell, der målefeilen er ukorrelert med teoretisk variabel X^* og dermed korrelert med observert variabel (Woolridge, kap 9-4.b). OLS inkonsistent, skjev mot null. Dette tilfellet er mest interessant, og bør kommentere hva som bestemmer størrelsen på skjevheten i OLS-estimatoren her og retningen på skjevheten. (signal to noise-ratio). Her er det også relevant å diskutere alternativ estimeringsmetode (Instrumentvariabelmetoden).

Hovedkonklusjon er at problemet med målefeil er avhengig av hvilke(n) variabel som inneholder målefeil.