

Eksamen på Datalogi-Økonomi Blok 1 2024

Økonometri A

1. november, 2024, kl. 10.00-22.00

Praktiske anvisninger til eksamen i Økonometri A

- Eksamenen kan besvares individuelt eller i grupper af maks. 3 studerende.
- Besvar alle spørgsmål og delspørsmål i opgave 1 til 4. Delopgaverne er adskilt med #.
- Rapporten skal forsynes med en forside og der skal udarbejdes en ansvarsfordeling, således de enkelte gruppemedlemmers bidrag til besvarelsen fremgår. Alle studerende i gruppen skal være ansvarlig for mindst et a-, b- og c-spørsmål. Dette gøres ved at anvende skabelonen forside.doc, som er tilgængelig på Digital Eksamen.
- Besvarelsen skal bestå af en samlet rapport med *relevante* tabeller og figurer.
- Alle sider i rapporten skal forsynes med sidetal og eksamensnumre.
- Der er *ikke noget sidekrav* ved denne eksamen, men prøv så vidt muligt at være koncis og undgå at præsentere mere materiale end nødvendigt. Undgå at eksempelvis at inkludere en omfattende liste af dummyvariable i tabellerne.
- Fontstørrelse skal sættes til mindst 12, linjeafstand sat til 1,5 (halvanden), og sidemarginer sat til mindst 2,5 cm. Figurer og tabeller skal nummereres.
- Eksamenen besvares ved at uploade rapporten inklusiv bilag til Digital Eksamen inden eksamenen er slut (husk at trykke Aflever). Dette gøres ved at følge instruktionerne beskrevet i følgende video link

- Rapporten uploades på Digital Eksamen som en pdf navngivet udelukkende med eksamensnumrene i gruppen forbundet med "_", f.eks. 12_34_56.pdf. Bemærk at der kun skal uploades én besvarelse for hver gruppe.
- Udover selve rapporten skal, der også uploades en forside (som beskrevet ovenfor) og et Python program eller Jupyter Notebook til Digital Eksamen.

Python-programmet skal kunne eksekveres uden fejl og generere resultater i samme rækkefølge, som de fremgår af rapporten.

- Hvis der mod forventning skulle være fejl eller uklarheder i eksamensopgaven eller problemer med at downloade eller indlæse data, kan I skrive en email direkte til Bertel Schjærning (bertel.schjærning@econ.ku.dk). I vil ikke altid få personligt svar, men hvis det er nødvendigt at præcisere eller rette noget af opgaveteksten vil I alle blive kontaktet på jeres KU-email via Absalon. I skal derfor sørge for at tjekke jeres email under eksamenen.
- Hvis Digital Eksamen er ramt af nedbrud, kan besvarelsen i nødstilfælde indsendes via et gruppe-medlems KU-email til uddannelse@diku.dk. Dette kræver dog en udførlig dokumentation af problemet inklusiv et skærmpoint. Fag og eksamensnumre skal fremgå af emailen.
- Det er ikke tilladt at kommunikere med andre grupper under eksamenen. Enhver form for kommunikation betragtes som eksamenssnyd og vil blive behandlet herefter.
- Dette eksamenssæt består af 15 sider i alt.

Pas på du ikke begår eksamenssnyd!

Det er for eksempel eksamenssnyd, hvis du ...

- Kopierer andres tekster uden at sætte citationstegn eller kildehenvise, så det ser ud, som om, det er din egen tekst.
- Bruger andres idéer eller tanker uden at kildehenvise, så det ser ud, som om det er din egen idé eller dine tanker.
- Genbruger dele af en opgave, som du tidligere har indleveret og fået en bestået karakter for uden at sætte citationsregn eller kildehenvise (selvplagie ring).
- Modtager hjælp fra andre i strid med gældende regler.

Eksamenssnyd sanktioneres altid med en advarsel og bortvisning fra prøven. I de fleste tilfælde bliver den studerende også bortvist fra universitetet i et semester. Du kan læse mere om reglerne vedrørende eksamenssnyd på Absalon under eksamensform og -regler [Link](#)

Introduktion til opgaven

Bilmarkedet spiller en central rolle i moderne økonomier som både en kilde til beskæftigelse, innovation og som en vigtig faktor i transportsektorens klimaafttryk. Dette marked er komplekst, idet producenter konkurrerer om at levere biler, der matcher forbrugernes præferencer, mens regeringer ofte regulerer markedet gennem skatter, afgifter og miljøstandarder. En afgørende udfordring i analyser af bilmarkedet er at forstå, hvordan forbrugerne reagerer på ændringer i priser og produktkarakteristika, og hvordan virksomheder fastsætter deres priser under hensyntagen til disse reaktioner. Ved at kombinere data, økonomisk teori med statistiske metoder kan vi opnå en dybere indsigt i disse mekanismer og kvantificere effekterne af forskellige politikker på bilmarkedet.

Vi tager udgangspunkt i en strukturel model udbud og efterspørgsel efter biler inspireret af Grieco, Murry and Yurukoglu (2024). Vi undersøger, hvordan forbrugere træffer valg mellem forskellige bilmodeller baseret på bilernes priser og egenskaber, og hvordan virksomhedernes prissætning kan bruges til at udlede virksomhedernes marginalomkostninger, under antagelsen om, at virksomhederne sætter deres priser optimalt. Denne strukturelle tilgang giver os mulighed for at estimere parametre, som gør det muligt at simulere effekterne af forskellige politikker på bilmarkedet, såsom skatter på hestekræfter eller ændrede tariffer.

Eksamenssættet består af 4 opgaver, hvor de første 3 opgaver fokuserer på empiriske analyser af bilmarkedet, mens og den sidste opgave omfatter en teoretisk og en simulationsbaseret Monte Carlo-analyse, hvor vi gennem simuleringer undersøger OLS- og IV-metodernes egenskaber under forskellige betingelser.

Model for udbud, efterspørgsel og ligevægt på bilmarkedet

Forbrugerne efterspørgsel efter biler

Vores analyse begynder med modelleringen af forbrugernes valg af biler. Forbrugerne maksimerer deres nytte ved valg mellem forskellige bilmodeller, $j = 0, 1, \dots, J$, herunder muligheden for ikke at købe en bil, $j = 0$. Vi antager, at forbrugerne vælger det produkt, j^* , der maksimerer deres nytte, givet prisen og produktkarakteristika for produkt j på tidspunkt t :

$$j_{it}^* = \arg \max_{j=0,1,\dots,J} U_{ijt},$$

hvor $j = 0$ repræsenterer outside good (ikke at købe en bil).

Vi antager, at forbrugernes præferencer kan beskrives ved en lineær nyttefunktion

$$U_{ijt} = \beta \mathbf{x}_{jt} - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt},$$

hvor \mathbf{x}_{jt} er observerbare karakteristika for produkt j i marked t (f.eks. hestekræfter, brændstofeffektivitet), p_{jt} er prisen for produkt j i marked t , ξ_{jt} er uobserverbare karakteristika, som kan være brandimage eller andre kvaliteter, der ikke er direkte målt, og ϵ_{ijt} er idiosynkratisk præference for forbruger i , antaget at følge en ekstremværdifordeling.

Antagelsen om, at ϵ_{ijt} følger en ekstremværdifordeling, leder til aggregerede markedsandele på logit-formen

$$s_{jt} = \frac{\exp(\beta \mathbf{x}_{jt} - \alpha p_{jt} + \xi_{jt})}{1 + \sum_{k=1}^J \exp(\beta \mathbf{x}_{kt} - \alpha p_{kt} + \xi_{kt})},$$

hvor s_{jt} er sandsynligheden for, at forbruger i vælger produkt j i marked t .

Ved at normalisere nytten af outside good til 0, kan vi skrive log-markedsandelen relativt til andelen af husholdninger uden bil:

$$\log\left(\frac{s_{jt}}{s_{0t}}\right) = \beta \mathbf{x}_{jt} - \alpha p_{jt} + \xi_{jt}. \quad (1)$$

Dette resulterer i en lineær regressionsmodel, hvor $\log(s_{jt}) - \log(s_{0t})$ er den afhængige variabel, og \mathbf{x}_{jt} og p_{jt} er observerbare forklarende, ξ er fejlleddet, og β og α er parametre vi ønsker at estimere.

Virksomhedernes prissætning og marginalomkostninger

Vi antager, at virksomhederne opererer i uafhængige markeder og sætter prisen p_{jt} , der maksimerer profit givet efterspørgselsfunktionen for produkt j i marked t

$$p_{jt} = \arg \max_p (p - mc_{jt}) s_{jt},$$

hvor mc_{jt} er marginalomkostninger for virksomhed j i marked t . Førsteordensbetingelserne for optimal prissætning er givet ved

$$s_{jt} + (p_{jt} - mc_{jt}) \frac{\partial s_{jt}}{\partial p_{jt}} = 0.$$

Denne betingelse giver den optimale prissætning som en funktion af marginalomkostninger og efterspørgslen:

$$p_{jt} = mc_{jt} - s_{jt} \left(\frac{\partial s_{jt}}{\partial p_{jt}} \right)^{-1} = mc_{jt} + \frac{1}{\alpha(1 - s_{jt})}. \quad (2)$$

hvor det sidste lighedstegn følger at differentiere efterspørgselsfunktionen i (1) mht. til prisen $\frac{\partial s_{jt}}{\partial p_{jt}} = -\alpha s_{jt}(1 - s_{jt})$. Vi antager, at marginalomkostningerne for virksomhed j i marked t kan beskrives som en lineær funktion af produktkarakteristika, \mathbf{x}_{jt} og udbudsspecifikke karakteristika, \mathbf{w}_{jt} :

$$\log(mc_{jt}) = \gamma_x \mathbf{x}_{jt} + \gamma_w \mathbf{w}_{jt} + \nu_{jt}, \quad (3)$$

hvor \mathbf{x}_{jt} er de observerbare produktkarakteristika som også påvirker nytten (f.eks. størrelse, kvalitet), \mathbf{w}_{jt} er udbudsspecifikke karakteristika (f.eks. råvarepriser, skatter, toldmure, lønomkostninger, mv), ν_{jt} er en stokastisk udbudsshock, og γ_x og γ_w er parametre, vi ønsker at estimere.

Ligevægt på bilmarkedet

Modellen består altså af tre ligninger, (1), (2), og (3), som introducerer en simultanitet mellem priser og markedsandele, som vi er nødt til at tage højde for, hvis vi ønsker at estimere modellens parametre ved at bruge data på priser, markedsandele, s_{jt} , produktkarakteristika, \mathbf{x}_{jt} , og udbudsspecifikke variable, \mathbf{w}_{jt} . Det virker umiddelbart som en vanskelig opgave, da både markedsandelene, s_{jt} og prisen p_{jt} er endogene fordi de er bestemt i en ligevægt. Det er den udfordring, vi står overfor i denne opgave. Heldigvis tar vi et skridt ad gangen.

Data

Til eksamen i Økonometri A er der adgang til datasættet `cars.csv` på Digital Eksamen. Der er kun et datasæt, som alle grupper skal bruge. Data-filen kan indlæses som `DataFrame` i Python med følgende kode

```
import numpy as np
df = pd.read_csv("cars.csv")
df.describe(percentiles=[]).T
```

Verificer, at data kan indlæses uden fejl. Datasættet indeholder variablerne beskrevet i Tabel 1.

Tabel 1: Variabelbeskrivelse

Variabel	Beskrivelse og Værdisæt
year	Kalenderår for observation. Interval, $t = 1980, 1981 \dots 2018$
product_id	Unik ID for make-model, $j = 0, \dots, J = 1066$
make	Bilproducent som Ford, Toyota, BMW, Volvo, etc.
model	Model af køretøj som BMW 3 Series, Chevrolet Suburban, Ford Mustang, etc.
number_households	Antal husholdninger i året, n_t
sales	Solgte enheder i et givent år
price	Anbefalet udsalgspris, p_{jt} , målt i \$1,000 (2015-priser)
height	Højde på køretøjet. Målt i tommer
footprint	Areal (længde \times bredde). Målt i kvadrattommer
weight	Vægt af køretøjet. Målt i pund
hp	Hestekræfter.
mpg	Brændstofeffektivitet målt i miles per gallon
yearsSinceDesign	Antal år siden design blev introduceret. Interval: 0-25
number_trims	Antal trimniveauer for modellen
releaseYear	Dummy for nyt modelår
RXR	real valutakurs
lux	Dummy for luksusbiler
sport	Dummy for sportsbiler
HEV	Dummy for hybridbiler
PHEV	Dummy for plug-in hybridbiler
EV	Dummy for elbiler
car	Dummy for personbiler
suv	Dummy for SUV'er
truck	Dummy for lastbiler
van	Dummy for varevogne
market_share	Markedsandel, $s_{jt} \in (0, 1)$
no_car	Andel af husholdninger uden bil, $s_{0t} \in (0, 1)$
dlogS	Logaritme af markedsandele relativt til andelen uden bil, $\log(s_{jt}) - \log(s_{0t})$

Note: Datasættet indeholder kun et udvalg af variablerne fra det oprindelige datasæt. Nogle variabler er omdøbt eller skaleret anderledes.

Data indeholder 9.694 observationer fordelt på årene 1980-2018. Enhed for observationer er (j, t) , dvs. produkt-år kombinationer (eller make-model-år kombinationer). Priser (price) er målt i \$1.000 (2015-priser). year og number_households varierer kun over tid. Dummyvariablene car, suv,

truck, og van er gensidigt udelukkende og summerer til 1. Dummyvariablene EV, HEV, og PHEV er også gensidigt udelukkende, men de summerer ikke til 1 (f.eks. mangler dummy for diesel- og benzinbiler). De resterende to dummier, sport og lux, er ikke gensidigt udelukkende og der er ingen dummy for referencekategorien.

Opgave 1 (25%)

(1A) Udfør en deskriptiv analyse af datasættet `data.csv`. Illustrér udviklingen over tid i priser, markedsandele og andre centrale attributter som hestekræfter (`hp`), brændstofeffektivitet (`mpg`), areal (`footprint`) og vægt (`weight`) for forskellige biltyper (f.eks. `car`, `suv`, `truck`, `van`). Undersøg sammenhængen mellem logaritmen af markedsandele relativt til andelen uden bil (`dlogS`) og bilpriser, samt mellem priser og fysiske attributter for bilerne. Kommentér kort på de observerede tendenser.

#

(1B) Estimér følgende lineære model baseret på efterspørgselsligningen i ligning (1), hvor den afhængige variabel er logaritmen af markedsandelen relativt til andelen uden bil $y_{jt} = \log(\frac{s_{jt}}{s_{0t}})$

$$y_{jt} = \beta \mathbf{x}_{jt} - \alpha p_{jt} + \delta_t + \xi_{jt}.$$

hvor \mathbf{x}_{jt} inkluderer et konstantled, produktkarakteristika som `hp`, `mpg`, `height`, `footprint`, `weight`, `number_trims`, `releaseYear`, og `yearsSinceDesign`, samt dummies for biltyper som `sport`, `EV`, `truck`, `suv` og `van`. Der kontrolleres for tidseffekten, δ_t , ved at inkludere et fuldt sæt af tids-dummies for kalenderår dannet ud fra variable `year`. Tidsdummies kan let genereres i Python ved at bruge `pd.get_dummies` funktionen i `pandas`. Eksempelvis:

```
# Dummy for 'year'
df = pd.get_dummies(df, columns=['year'], drop_first=True)
```

Estimer flere specifikationer. Rapportér parameterestimer, standardfejl, TSS, ESS, RSS og R^2 for modellerne. Fortolk estimatet for α og de vigtigste β -koefficienter (effekten af produktkarakteristika og biltyper). Diskutér, hvordan kontrol for forskellige variable påvirker α . Estimér modellen med produktkarakteristika i deres oprindelige skala (levels) og med log-transformationer af de kontinuerte variable `hp`, `mpg`, `height`, `footprint`, `weight` og `number_trims`. Sammenlign specifikationerne mht. R^2 , og forklar forskellene i estimerne. Vurder, hvilken model bedst beskriver data.

#

(1C) Udfør et test af $H_0 : \alpha = 0$ mod det tosidede alternativ $H_A : \alpha \neq 0$. Opstil den relevant test statistik, redegør for de nødvendige antagelser bag testet og konkluder på testresultatet. Diskutér, hvordan testresultatet påvirkes af, om der inkluderes produktkarakteristika, biltyper og tidseffekter i modellen, samt hvorvidt resultatet er følsomt over for anvendelsen af heteroskedasticitetsrobuste standardfejl.

Opgave 2 (25%)

(2A) Udvid OLS-modellen fra 1B ved at inkludere bilmærke-specifikke fixed effects baset på variable `make`. Vi dekomponerer efterspørgsels chokket som $\xi_{jt} = \bar{\xi}_m + \tilde{\xi}_{jt}$, hvor $\bar{\xi}_m$ er bilmærke-specifikke effekter, som er konstante over tid for hvert bilmærke, og $\tilde{\xi}_{jt}$ er den resterende del, der varierer både med bilmodel og tid. Vi har altså:

$$y_{jt} = \beta \mathbf{x}_{jt} - \alpha p_{jt} + \delta_t + \bar{\xi}_m + \tilde{\xi}_{jt}, \quad (4)$$

Estimér først modellen ved at inkludere `make`-specifikke dummies, og estimer derefter modellen ved hjælp af within-transformationen:

$$y_{jt} - \bar{y}_m = \beta (x_{jt} - \bar{x}_m) - \alpha (p_{jt} - \bar{p}_m) + (\delta_t - \bar{\delta}) + (\tilde{\xi}_{jt} - \bar{\xi}_m), \quad (5)$$

hvor $\bar{\cdot}$ angiver gennemsnittet inden for hvert bilmærke m baseret på variabelen `make`. For eksempel beregnes $\bar{x}_m = \frac{1}{T_m} \sum_{j \in m} x_{jt}$, hvor T_m er antallet af observationer for bilmærke m . De øvrige gennemsnit, såsom \bar{p}_m , er defineret på tilsvarende måde. Within-transformationen kan let implementeres i Python ved at bruge `groupby` og `transform` funktioner i `pandas`. Eksempelvis:

```
def within(df, group_var, vars_to_transform):
    # Within-group-var-transformation for alle variable i vars_to_transform
    df_within = df.copy()
    for var in vars_to_transform:
        df_within[var] = df[var] - df.groupby(group_var)[var].transform('mean')
    return df_within

# Eksempel på brug: Anvend within-make-transformation på udvalgte variable
df_within = within(df, 'make', ['dlogS', 'price', 'hp'])
```

Sammenlign parameterestimererne for de to modeller. Er de ens? Forklar, hvorfor konstantleddet er nul i modellen med within-transformerede variable, og hvorfor $\bar{\xi}_m$ forsvinder i modellen efter within-transformationen. Sammenlign også TSS, ESS, RSS og R^2 for de to modeller. Diskuter betydningen af eventuelle forskelle i goodness-of-fit mål. Forklar, hvorfor R^2 er højere i modellen med `make`-dummies sammenlignet med den within-transformerede model.

#

(2B) Sammenlign med specifikationen i 1B, og diskuter nødvendigheden af at inkludere `make`-specifikke fixed effects for at sikre konsistente estimater af α og β . Ifølge (2) er prisen p_{jt} bestemt af marginalomkostningerne, mc_{jt} , og efterspørgslen, s_{jt} . Kan vi ud fra den strukturelle model forvente, at de bilmærke-specifikke effekter, $\bar{\xi}_m$, er positivt eller negativt korreleret med prisen p_{jt} ? Hvilken retning

forventer du, at bias i estimatet af α vil tage, hvis make-effekterne udelades fra modellen? Er det konsistent med det, du finder?

#

(2C) Selvom inkluderingen af make-specifikke fixed effects eliminerer $\bar{\xi}_m$ fra modellen, kan den resterende tids- og modelvarierende komponent, $\tilde{\xi}_{jt}$, stadig være korreleret med prisen p_{jt} . For at adressere dette anvender Grieco, Murry and Yurukoglu (2024) den reale valutakurs, RXR, som instrument. RXR påvirker især udenlandske bilproducenters marginalomkostninger og indgår derfor i de udbudsspecifikke faktorer w_{jt} i ligning (3), som igen påvirker udbudsprisen p_{jt} gennem ligning (2).

Bemærk, at selvom den funktionelle form for den reducerede form for pris funktionen p_{jt} er ikke lineær i de eksogene variable, er en korrekt specificering af denne ikke et krav for at benytte RXR som instrument. Vi anvender derfor de flg. lineære approximation af den reducerede form for pris funktionen:

$$p_{jt} = \gamma_0 + \gamma_1 \text{RXR}_{jt} + \gamma_2 \mathbf{x}_{jt} + \kappa_t + \phi_m + v_{jt}, \quad (6)$$

hvor \mathbf{x}_{jt} er de eksogene produktkarakteristika fra efterspørgselsligningen, og κ_t og ϕ_m er henholdsvis tids- og make-dummies, der kontrollerer for tids- og mærkespecifikke effekter.

Estimer først (6) as the "first stage regression", og test, om RXR er en signifikant prædiktør for prisen ved hjælp af et t-test af γ_1 . Estimerer herefter den strukturelle efterspørgselsligning (4) ved hjælp af 2SLS, hvor RXR anvendes som instrument for den endogene pris p_{jt} . Sammenlign 2SLS-estimatet af α med OLS-estimatet og vurder, om 2SLS-estimatet stemmer overens med din forventning om endogenitetsbias i OLS-estimatet. Argumentér også med henvisning til den teoretiske model, hvorfor RXR er et gyldigt instrument for prisen p_{jt} .

Opgave 3 (25%)

(3A) Gandhi and Houde (2019) foreslår at anvende instrumenter baseret på forskelle i produktkarakteristika. Disse instrumenter måler antallet af konkurrerende biler, der har karakteristika tæt på de undersøgte biler, og er særligt relevante i markeder med differentierede produkter, hvor priserne tilpasses konkurrencen. Vi vil derfor afprøve sådanne instrumenter:

$$z_{jt}^{(g)} = \sum_{k \in J} 1 \left(|x_{jt}^{(g)} - x_{kt}^{(g)}| < sd_{gt} \right), \quad (7)$$

hvor $x_{jt}^{(g)}$ er den g 'te normaliserede karakteristik af bil j på tidspunkt t , og sd_{gt} er standardafvigelsen for den pågældende karakteristik beregnet for hvert år t . Dette sikrer, at instrumenterne afspejler variationen i konkurrencen blandt biler inden for det pågældende års marked. Vi kan implementere disse instrumenter i Python ved at bruge følgende funktion:

```
def gh_iv(df, characteristics):
    df_iv = df.copy() # Opret en kopi for at undgå ændringer i originalen

    for char in characteristics:
        # Normaliser karakteristikken inden for hvert år
        df_iv[char] = df_iv.groupby('year')[char].transform(lambda x: x /
            ↪ x.std())

        # Beregn antallet af biler tæt på hver bils karakteristika
        df_iv[f'gh_iv_{char}'] = df_iv.groupby('year').apply(
            lambda group: group[char].apply(lambda x: ((group[char] - x).abs()
            ↪ < group[char].std()).sum())
        ).reset_index(level=0, drop=True)

    return df_iv[[f'gh_iv_{char}' for char in characteristics]] # Returner kun
    ↪ instrumentkolonnerne
```

Diskuter, hvorfor disse instrumenter er relevante og forventes at være gyldige til at identificere parameteren α . Estimér efterspørgselsligningen ved hjælp af 2SLS med de udvidede instrumenter, og udfør en overidentifikationstest for at vurdere instrumenternes validitet. Sammenlign resultatet med tidligere estimer og diskuter eventuelle forskelle. Overvej fordele og ulemper ved at bruge disse instrumenter sammenlignet med RXR.

#

(3B) Vi skal nu bruge de estimerede efterspørgselsligninger til at beregne markedsandele, marginalomkostninger og priselasticiteter for forskellige bilmodeller og over tid. Givet et estimat for α , de observerede priser, og markedsandele, kan vi beregne marginalomkostningerne som:

$$p_{jt} = mc_{jt} + \frac{1}{\alpha(1 - s_{jt})} \quad (8)$$

$$\Leftrightarrow mc_{jt} = p_{jt} - \frac{1}{\alpha(1 - s_{jt})}. \quad (9)$$

Egenpriselasticiteten kan beregnes som:

$$\epsilon_{jt} = \frac{\partial s_{jt}}{\partial p_{jt}} \cdot \frac{p_{jt}}{s_{jt}} = -\alpha \cdot p_{jt} \cdot (1 - s_{jt}).$$

Markup kan beregnes som:

$$\text{Markup}_{jt} = \frac{p_{jt} - mc_{jt}}{p_{jt}}$$

Beregn marginalomkostningerne, markups og priselasticiteter for alle observationer i datasættet baseret på både OLS og IV-estimererne opnået i 2C, $\hat{\alpha}^{OLS}$ og $\hat{\alpha}^{IV}$. Har du ikke estimeret α , så brug værdierne $\hat{\alpha}^{OLS} = 0.04$ og $\hat{\alpha}^{IV} = 0.15$.

Udfør en diskriptiv analyse for at undersøge, hvordan marginalomkostninger, priselasticiteter og markups varierer over tid og mellem forskellige bil-typer. Illustrer også fordelingen af markedsandele, priser, marginalomkostninger og markups for forskellige biltyper. Brug gerne visualiseringer.

Reflektér over, hvor afhængige vi er af modeller og antagelser, når vi estimerer marginalomkostninger. Diskutér også, hvilke konsekvenser det kan have, hvis modellen er fejlspecificeret, eller instrumenterne ikke hvis ikke er valide eller stærke nok til at korrigere for prisendogenitet.

#

(3C) Brug de predikterede værdier af marginalomkostningerne fra 3B baseret på IV-estimererne for α til at estimere marginalomkostningsfunktionen (3) med bilmærke- og tidsspecifikke effekter:

$$\log(mc_{jt}) = \gamma_x \mathbf{x}_{jt} + \gamma_{RXR} RXR_{jt} + \phi_m + \kappa_t + \nu_{jt} \quad (10)$$

Fortolk parameterestimererne, γ_x og γ_{RXR} , og diskutér hvordan marginalomkostningerne varierer med produktkarakteristika og RXR, og om marginalomkostningerne er steget eller faldet over tid, når vi kontrollerer for produktkarakteristika, RXR, og mærkespecifikke effekter?

#

Opgave 4 (25%)

(4A) Betragt en simpel version af den strukturelle model for efterspørgslen efter biler:

$$\log(s_i) = \beta_0 - \alpha p_i + \xi_i, \quad i = 1, \dots, n \quad (11)$$

hvor ξ_i repræsenterer uobserverede efterspørgselsfaktorer, og hvor vi antager, at prisen p_i afhænger lineært af z_i og et udbudsshock ν_i :

$$p_i = \gamma_0 + \gamma_1 z_i + \nu_i, \quad (12)$$

hvor z_i er et eksogent instrument, der påvirker prisen, men ikke direkte $\log(s_i)$, og ν_i er et udbudssho-ck, som kan være korreleret med ξ_i .

Indsæt udtrykket for p_i i ligningen for $\log(s_i)$, og udled den reducerede form for $\log(s_i)$ som en funktion af z_i . Diskutér, om det er muligt at identificere α udelukkende baseret på den reducerede form for $\log(s_i)$. Hvis ikke, hvad er årsagen til dette, og hvordan kan vi løse identifikationsproblemet?

#

(4B) Antag, at ξ_i kan skrives som en lineær funktion af ν_i og et uafhængigt stød η_i :

$$\xi_i = \rho\nu_i + \eta_i, \quad (13)$$

hvor η_i er uafhængig af ν_i . Lad $\alpha = 0.1$, $\beta_0 = 0$, $\gamma_0 = 1$, $\gamma_1 = 0.5$, og antag, at $z_i \sim N(0, 1)$, $\nu_i \sim N(0, 1)$, og $\eta_i \sim N(0, 1)$.

Udled biasen for OLS, når $\log(s_i)$ regresseres mod prisen og en konstant, og vis, at IV-estimatoren er konsistent for priskoefficienten $-\alpha$. Hvilken retning forventer du, at bias i OLS-estimatet vil tage, hvis ρ er positiv? Hvordan påvirker styrken af instrumentet, z_i , bias og varians af IV-estimatoren?

#

(4C) Simuler data baseret på ovenstående model givet ved (11), (12) og (13), og udfør et Monte Carlo-eksperiment med $n = 500$ observationer under følgende scenarier:

- $\rho = 0$ (ingen korrelation mellem ξ_i og ν_i).
- $\rho = 0.5$ (moderat korrelation).
- $\rho = 0.9$ (stærk korrelation).

Estimer α ved OLS og IV for hvert scenarie, hvor z_i bruges som instrument for p_i . Beregn og sammenlign bias og varians af OLS- og IV-estimatorerne. Diskutér resultaterne med fokus på, hvordan korrelationen ρ påvirker estimatorernes konsistens og bias. Undersøg også, hvordan styrken af instrumentet påvirker bias og varians af IV-estimatoren. Sammenhold med resultaterne fra 4B.

#

Litteratur

GANDHI, A. AND J.-F. HOUDE (2019): “Measuring substitution patterns in differentiated-products industries,” *NBER Working paper*, (w26375).

GRIECO, P. L., C. MURRY AND A. YURUKOGLU (2024): “The evolution of market power in the us automobile industry,” *The Quarterly Journal of Economics*, 139(2), 1201–1253.