

# Introduktion

Økonometri A

---

Bertel Schjerning

September 2, 2024

# Program

Hvad er økonometri? (W1.1-W1.2)

Forskellige typer af data (W1.3-W1.4)

Eksempel: Structural Model for Consumer Demand

Fagbeskrivelse

Praktisk information

**Hvad er økonometri?**

---

# Hvad er økonometri?

Økonometri kombinerer tre grundelementer:

- **Økonomiske Modeller:** Teoretiske relationer, der beskriver økonomisk adfærd.
- **Data:** Økonomiske data indsamlet fra virkeligheden.
- **Statistiske Metoder:** Værktøjer til at analysere data og teste hypoteser.

# Hvad er økonometri?

## Økonometri kombinerer tre grundelementer:

- **Økonomiske Modeller:** Teoretiske relationer, der beskriver økonomisk adfærd.
- **Data:** Økonomiske data indsamlet fra virkeligheden.
- **Statistiske Metoder:** Værktøjer til at analysere data og teste hypoteser.

## Formålet med økonometri

- Kvantificere økonomiske sammenhænge og teste økonomiske teorier.
- Estimere parametre i økonomiske modeller ved brug af data.
- Evaluere effekten af økonomiske politiktiltag og interventioner.

## Den simple lineære regressionsmodel:

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

- $y_i$ : Den afhængige variabel (f.eks. løn).
- $x_i$ : Den uafhængige variabel (f.eks. uddannelse).
- $\beta_0, \beta_1$ : Parametre, der skal estimeres.
- $u_i$ : Fejlleddet, som fanger alle andre faktorer, der påvirker  $y_i$ .

Modellen antager en lineær sammenhæng mellem  $x_i$  og  $y_i$ .

# Eksempel på Lineær Model

## Løn og Uddannelse:

$$\text{Løn}_i = \beta_0 + \beta_1 \text{Uddannelse}_i + u_i$$

- $\beta_1$  repræsenterer den forventede stigning i løn for hvert ekstra års uddannelse.
- Ved at estimere  $\beta_1$  kan vi kvantificere uddannelsens effekt på lønnen.

Resultaterne fra denne model kan hjælpe med at informere politikker, der sigter mod at øge uddannelsesniveaue.

# Teori, data og statistiske metoder

- **Økonomisk teori:**

- Giver en strukturel beskrivelse af økonomisk adfærd og relationer mellem økonomiske variable.
- Bestemmer strukturen af den lineære model, inkl. hvilke variable og parametre der skal inkluderes.
- Forudsiger, hvordan ændringer i  $x$  (den uafhængige variabel) påvirker  $y$  (den afhængige variabel).

- **Data:**

- Data består af observerede værdier for  $x$  og  $y$  og bruges til at kvantificere den økonomiske relation.
- Data bruges til at opnå estimater af modellens parametre ( $\hat{\beta}_0, \hat{\beta}_1$ ).

- **Statistiske metoder:**

- Anvendes til at vælge passende estimatorer for  $\beta_0$  og  $\beta_1$ .
- Gør det muligt at vurdere estimatorernes egenskaber, såsom middelrethed, efficiens og varians.
- Værktøjer til evaluering af modellens pålidelighed og til at teste økonomiske hypoteser.



# Hvad er økonometri? Eksempler

Får ejendomsmæglere en højere pris, når de sælger deres egne huse end når de sælger klienternes huse?<sup>1</sup>

- Når ejendomsmæglere sælger et hus for en klient får mægleren en fast andel af salgsprisen.
- Når han sælger sit eget hus får han hele salgsprisen.
- Mægleren har således et større incitament til at forsøge at få en ekstra høj pris for sit eget hus sammenlignet med klienternes huse.
- Data for mere end 100,000 hussalg i Illinois, USA, og af dem var 3,300 huse ejet af mægleren selv.
- Økonometrisk analyse: Huse ejet af mægleren selv blev solgt for 3.7% mere end huse som ikke var ejet af mægleren.

<sup>1</sup> Steven D. Levitt and Chad Syverson (2005): "Market Distortions when Agents are Better Informed: The Value of Information in Real Estate Transactions", NBER WP no. 11053.

# Hvad er økonometri? Eksempler

Vil mere politi på gaderne give mindre kriminalitet?<sup>2</sup>

- I følge Gary Becker<sup>3</sup> burde mere politi give mindre kriminalitet fordi sandsynligheden for at blive fanget er større.
- Data for Buenos Aires, Argentina: Antallet af biltyverier for 876 boligblokke (boligkareer) for ni måneder april-december i 1994 og information om der er politiovervågning af blokken.
- Økonometrisk analyse: Er der færre biltyverier i de boligblokke med politiovervågning?

---

<sup>2</sup> Di Tella and Schargrodsky. 2004. "Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack." *American Economic Review*, 94 (1):115-133

<sup>3</sup> Becker, Gary S. "Crime and punishment: An economic approach." *The economic dimensions of crime*. Palgrave Macmillan, London, 1968. 13-68.

## Hvad er økonometri? Eksempler

Hvordan ændrer folk deres forbrug, når de får mere indkomst?

- Engelkurver viser sammenhængen mellem budgetandelen anvendt på forskellige goder og indkomsten (Mikro A).
- Engelkurver anvendes ved beregning af velfærdseffekterne af ændrede indkomster, fattigdomsgrænser samt skattepolitik.
- Økonometrisk analyse: Hvordan påvirkes budgetandelen (på fx mad) af ændringer i indkomsten? Hvilke goder er luksusgoder og hvilke er nødvendighedsgoder?

Opgavesæt 1 og 2 ved holdundervisningen handler om at estimere Engelkurver.

# Hvad er økonometri? Eksempler

Eksempler på empiriske spørgsmål fra Makroøkonomi

- Hvorfor har nogle lande højere vækst?
- Hvorfor er nogle lande rige?
- Hvordan afhænger væksten af de initiale betingelser?
- Hvordan påvirker  $K/L$  forholdet BNP?

Opgavesæt 5 ved holdundervisningen handler om spørgsmål fra Makroøkonomi.

## Forskellige typer af data

---

Vi kan skelne mellem: Hvilken **struktur** har data?

- Tværsnitsdata (cross-section)
- Tidsseriedata (time series)
- Gentaget Tværsnitsdata (pooled/repeated cross section)
- Paneldata (panel data)

Og hvordan er data **skabt**?

- Eksperimentel data
- Ikke-Eksperimentel/Observeret data

# Datastruktur: Tværsnit

Tværsnitsdata er data for forskellige enheder (individer, firmaer, osv.) observeret på “samme” tidspunkt:

$$(x_i, y_i) : \quad i = 1, 2, \dots, n$$

Variationen i data er på tværs af enheder:

- Husstandes forbrug og indkomst i en enkelte period (fx 14 dage)
- Indviders lønninger og uddannelse målt i et givent år
- Kriminalitet i forskellige bydele i et givent år.

Centrale egenskab ved tværsnitsdata:

- Rangeringen af observationerne er ligegyldig.
- Observationerne er typisk uafhængige af hinanden

# Datastruktur: Tidsserie

Tidsseriedata er data for samme enhed (fx lande) observeret på forskellige tidspunkter:

$$(x_t, y_t) : \quad t = 1, 2, \dots, T$$

Variationen i data er over tid:

- Årlig BNP i Danmark over de sidste 30 år
- Dagsprisen på Bitcoin
- Renten på et 30-årigt realkreditlån

Centrale egenskab ved tværsnitsdata:

- Rangeringen af observationerne er vigtig.
- Observationerne næsten altid afhængige af hinanden



# Datastruktur: Panel

Paneldata indeholder både tværsnits- og tidsvariation:

$$(x_{it}, y_{it}) : \quad i = 1, 2, \dots, n \text{ og } t = 1, 2, \dots, T$$

Eksempler:

- Udviklingen i individuelle lønninger over tid .
- Efterspørgsel efter forskellige varer på forskellige markeder over tid.
- Udviklingen i aktiekursen for alle aktier i C25 indekset.

Centrale egenskaber ved paneldata:

- Kombinerer egenskaberne ved tværsnits- og tidseriedata.

# Eksperimentel vs. ikke-eksperimentel data

Som nævnt er vi typisk interesseret i **kausale effekter**.

- Dvs. effekten af at ændre en variable (fx mængden af gødning) på et givent outcome, **alt andet lige**

Guldstandarden for kausale studier er **randomiserede eksperimenter**

- En tilfældigt udvalgt gruppe udsættes for en særlig “behandling”, mens en anden gruppe ikke gør.
- Da grupperne er tilfældigt udvalgt vil der ikke være systematiske forskellene mellem dem bortset fra behandlingen.
- Forskellen i outcome mellem grupperne vil derfor kunne tilskrives behandlingen alene (alt andet er lige).

Randomiserede eksperimenter anvendes fx hyppigt i test af ny medicin.

# Eksperimentel vs. ikke-eksperimentel data

I samfundvidenskab er det ikke altid vi kan bruge eksperimentelle data.

- **Etiske overvejelser:** Det er ikke etisk forsvarligt at randomisere fx uddannelse.
- **Praktiske begrænsninger:** Det kan være meget dyrt at gennemføre eksperimenter, fx virkningen af skatteændringer.

Randomiserede eksperimenter bliver dog mere og mere udbredt i økonomi. Fx i

- Udviklingsøkonomi<sup>4</sup>
- Virksomhedsøkonomi<sup>5</sup> m.fl.

---

<sup>4</sup> Banerjee, Duflo, and Kremer. The influence of randomized controlled trials on development economics research and on development policy. The State of Economics.

<sup>5</sup> Lewis and Rao. The unfavorable economics of measuring the returns to advertising. The Quarterly Journal of Economics 130.4 (2015): 1941-1973.

# Eksperimentel vs. ikke-eksperimentel data

I fravær af eksperimenter bruger vi ikke-eksperimentel data.

- Dvs. data, hvor folk selv træffer de valg, vi er interesseret i at kende virkningen af.

Central spørgsmål ved ikke-eksperimentel data:

## Hvor kommer variationen fra?

- Hvorfor får nogle marker mere gødning?
- Hvorfor ansætter nogle byer flere betjente?

Kan vi forvente at alt andet er lige for folk, som har truffet forskellige valg?

## Eksperimentel vs. ikke-eksperimentel data

Eksempel 1: Bliver man sundere af at bo tæt på natur?

- **Randomiseret eksperiment:** Udvælg to tilfældige grupper. Giv den ene gruppe en bolig tæt på natur, mens den anden gruppe får en bolig langt væk fra natur. Sammenlign sundhedstilstanden for de to grupper.
- **Ikke-eksperimentelle data:** Data på sundhed for forskellige mennesker, hvoraf nogle bor tæt på natur. Sammenligner sundheden for dem som bor tæt på natur med dem som ikke gør.
- **Problem:** Måske bor sunde mennesker tættere på natur, fx pga højere indkomst.

## Eksperimentel vs. ikke-eksperimentel data

Eksempel 2: Får ejendomsmæglere en højere pris, når de sælger deres egne huse end når de sælger klienternes huse?

- **Randomiseret eksperiment:** Giv tilfældigt udvalgte huse til to grupper af mæglere, hvor den ene gruppe får alt profitten selv, mens den anden gruppe kun får en andel af profitten. Sammenlign salgspriserne for de to grupper.
- **Ikke-eksperimentelle data:** Data på huse, hvoraf nogle ejes af mægler. Sammenligner salgpris for huse ejet af mægler med huse som er ejet af en klient.
- **Problem:** Måske ejer mæglerne huse, som er mere attraktive.

## Eksperimentel vs. ikke-eksperimentel data

Eksempel 3: Vil mere politi på gaderne give mindre kriminalitet?

- **Randomiseret eksperiment:** Udvælg boligblokke tilfældigt og indsæt ekstra politi der. Derefter vil man kunne sammenligne antal af biltyverier for blokke med ekstra politi med blokke uden ekstra politi.
- **Ikke-eksperimentelle data:** Data på boligblokke, hvor der nogle steder er ekstra politi. Sammenligner antal biltyverier, hvor der er ekstra politi
- **Problem:** Man kan forestille sig at politiet er mere tilstede i områder med meget kriminalitet.

## Eksperimentel vs. ikke-eksperimentel data

Eksempel 4: Hvordan ændrer forbruget sig ved ændringer i indkomsten?

- **Randomiseret eksperiment:** Randomiser folk indkomst og sammenlign deres forbrugsmønstre.
- **Ikke-eksperimentelle data:** Data på indkomst og forbrug, hvordan nogen personer har høj indkomst og nogen har lav. Sammenlign personer høj indkomst med personer med lav.
- **Problem:** Vil folk med høj indkomst have samme forbrugsmønstre som folk med lav indkomst, hvis de havde samme indkomst eller er der forskelle i præferencer?



# Eksperimentel vs. ikke-eksperimentel data

Betyder det at ikke-eksperimentel data er ubrugeligt? Nej.

Flere strategier:

- **Kontrollerer** for andre forhold. Fx (observerede) kvalitetsforskelle mellem ejendomsmæglernes egne huse og deres klienter.
- **Isolerer** variation i data, som vi tror er uafhængig af andre (relevante) forhold. Fx indsættelse af politi ved synagoer efter et angreb på jøder.
- **Strukturel økonometri**: Anvender økonomisk teori til at specificere internt konsistente modeller, som kan bruges til at identificere kausale effekter og skabe ekstern validitet.

Disse strategier kommer vi til at arbejde med i dette fag.

# Forskellige økonometriske tilgange

- **Strukturel Økonometri:** Modeller forankret i økonomisk teori.
  - + Muliggør kontra-faktiske analyser og policy-eksperimenter.
  - - Kræver antagelser om modellens struktur.
- **Reduceret Form:** Modeller udledt fra strukturelle modeller ved at eliminere de endogene variabler.
  - + Mindre afhængig af specifikke modelantagelser.
  - - Kan ikke altid tolkes kausalt.
- **Kausal inferens:** Fokus på at identificere årsagssammenhænge, ofte ved hjælp af eksperimentelle eller semi-eksperimentelle metoder.
  - + Klar kausal fortolkning.
  - - Begrænset generaliserbarhed, praktiske udfordringer.
- **Beskrivende og Prædiktive Modeller:** Bruges til at beskrive data eller forudsige fremtidige observationer.
  - + Nyttig til at forstå komplekse data.
  - - Kan være svært at generalisere til andre sammenhænge.

## **Eksempel: Structural Model for Consumer Demand**

---

## Nyttefunktion for Produkt $j$ i Marked $m$ :

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

- $\mathbf{x}_{mj}$ : Observerbare karakteristika for produkt  $j$  i marked  $m$  (f.eks. størrelse, kvalitet).
- $p_{mj}$ : Pris for produkt  $j$  i marked  $m$ .
- $\xi_{mj}$ : Uobserverbare karakteristika (f.eks. brand, placering).
- $\epsilon_{mj}$ : Idiosynkratisk fejllid, der følger en ekstremværdi fordeling (Gumbel).

**Efterspørgselsfunktion for Produkt  $j$  i Marked  $m$ :**

$$S_{mj} = \frac{\exp(\beta \mathbf{x}_{mj} - \alpha p_{mj} + \xi_{mj})}{\sum_{k=1}^J \exp(\beta \mathbf{x}_{mk} - \alpha p_{mk} + \xi_{mk})}$$

**Log-Lineariseret Efterspørgsel:**

$$\log(S_{mj}) - \log(S_{m0}) = \beta \mathbf{x}_{mj} - \alpha p_{mj} + \xi_{mj}$$

**Hvordan Annulleres Log-Summen?**

- Ved at trække log-markedsandelen for outside-alternativet fra, elimineres summen i nævneren.
- Dette giver en lineær model i log-markedsandele.

## Endogenitetsproblemer:

- Prisen  $p_{mj}$  kan være korreleret med uobserverbare karakteristika  $\xi_{mj}$ .
- Eksempel: Produkter med højere uobserverbar kvalitet ( $\xi_{mj}$ ) kan have højere priser, hvilket skaber en bias i estimerterne.

## Mulige Instrumenter for Pris:

- **Omkostnings-shiftere:** Produktionsomkostninger, der påvirker prisen, men ikke direkte efterspørgslen.
- **Markedsvariationer:** Exogene variationer i udbud eller skatter, der påvirker priser på tværs af markeder, men ikke direkte efterspørgslen.
- **Exogene shock:** Naturlige eksperimenter eller policy-ændringer, der påvirker priser uden at påvirke forbrugerpræferencer direkte.

# Fordele ved Logit-demand Modellen med Instrumentvariable

## Fordele:

- Tillader estimation af elasticiteter og følsomhed over for prisændringer.
- Muliggør kontrafaktiske analyser, f.eks. hvordan ændringer i priser eller karakteristika påvirker markedsandele.
- Instrumentvariable (IV) kan korrigere for endogenitet, hvilket giver konsistente estimater.

## Ulemper:

- Kræver stærke instrumenter, som er korreleret med prisen, men ikke med uobserverede karakteristika.
- Kan være udfordrende at identificere passende instrumenter i praksis.

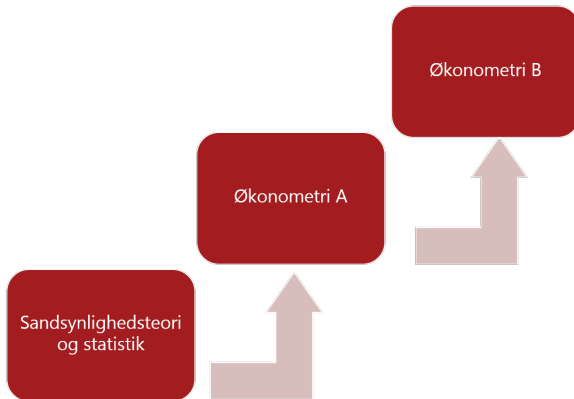
# Fagbeskrivelse

---



# Fagrækken i økonometri

Det andet fag i en række af tre økonometrifag: Grundlæggende statistik og sandsynlighedsregning, Økonometri A og Økonometri B



Økonometri A: **Estimation og inferens i den lineære model på tværsnit- og paneldata**

## Målbeskrivelse: Viden

- Redegøre for den multiple lineære regressionsmodel.
- Beskrive og forklare de mest centrale antagelser i regressionsanalyse, samt hvordan disse påvirker de estimerede parameterverdier.
- Diskutere og forklare om parameterestimerer kan tillægges en kausal fortolkning.
- Forklare og udføre statistisk inferens for den lineære regressionsmodel.
- Anvende økonometriske kriterier til at vælge det mest foretrukne sæt af parameterestimerer blandt flere muligheder.
- Redegøre for fortolkningen af parametre for kontinuerte, diskrete og transformerede variable i en regressionsmodeller.

## Målbeskrivelse: Færdigheder

- Udføre deskriptive analyser af datasæt med henblik på anvendelse i regressionsanalyse.
- Udlede simple estimatorer samt karakterisere deres statistiske egenskaber i form af middelværdi, varians, konsistent og effektivitet.
- Udføre og implementere estimationsmetoder (OLS, WLS, GLS, IV, DD, FD, FE, RE)
- Beregne t-test, F-tests, LM-tests, og Wald-test for hypotesetest i den lineære regressionsmodel med homoskedastiske og heteroskedastiske fejl.
- Udføre test for misspecifikation (heteroskedasticitet, funktionel form, eksogenitet m.fl.) og redegøre for deres fortolkning.
- Anvende simulationseksperimenter til at illustrere og efterprøve egenskaber for statistiske estimatorer og test.

## Målbeskrivelse: Kompetencer

- Planlægge og gennemføre en empirisk analyse af et selvstændigt udvalgt emne med udgangspunkt i de gennemgåede økonometriske metoder.
- Anvende økonometriske metoder til at vurdere og vælge mellem økonomiske teorier og på den måde skabe selvstændig og evidensbaseret viden.
- Anvende givne parameterestimer i en konkret beregning på en økonomisk problemstilling og redegøre for resultaterne.

Eksamensopgaver sigter på bredt at teste ovenstående viden, færdigheder og kompetencer.

## **Praktisk information**

---

**GitHub** kursushjemmesiden:

<https://github.com/bschjerning/EconometricsA>

**Absalon** kursushjemmesiden:

<https://absalon.ku.dk/courses/77309>

Stort set alt materiale til kurset bliver samlet på GitHub

## Pensum:

- Hovedtekst: J.M. Wooldridge: Introductory Econometrics: A Modern Approach, 7th Edition, 2019.
- Kapitel 1-9, 13-15, 19 og Appendices A-E.
- Noter:
  - Simulation Experiments in Econometrics (Rasmus Jørgensen, 2015)
  - Instrumental Variables Estimation (Søren Leth-Petersen, 2016)

## Forelæsninger

- Overblik og motivation
- Økonometrisk teori: udledninger og beviser
- Eksempler og Python illustrationer

## Holdundervisning

- Opgavesæt
- Praktisk erfaring med at lave analyser, brug af Python, teoriopgaver, fortolkning af resultater

**3 Obligatoriske opgaver:** Skal bestås for at komme til eksamen.

**Eksamen:** 12-timers tag-hjem-eksamen, grupper af max 3 studerende.

## Tidspunkter:

- Mandage kl. 10.15 – 12.00
- Tirsdage kl. 13.15 – 15.00
- Fredage kl. 10.15 – 12.00

## Indhold:

- Følger tekstbogen tæt - suppleret med få forelæsningsnoter.
- Jupyter notebooks bruges til illustrative eksempler (Data, Python Code, Output).
- Plan for forelæsninger er på GitHub kursushjemmeside  
[https://github.com/bschjerning/EconometricsA/blob/main/Course\\_Schedule.md](https://github.com/bschjerning/EconometricsA/blob/main/Course_Schedule.md)



Tidspunkter:

- Mandage kl. 08.15 – 10.00
- Tirsdage kl. 15.15 – 17.00
- Fredage kl. 08.15 – 10.00

Indhold: 8 opgavesæt

- Arbejde med rigtige data som skal analyseres
- Python programmering (både Python procedurer og programmer I selv skal skrive)
- Simulationsopgaver
- Teoriopgaver
- Fortolkning- skrive konklusioner og opsummeringer

## Obligatoriske opgaver

- Der er 3 obligatoriske afleveringsopgaver i løbet af blokken
- Opgaverne skal godkendes for at kunne indstilles til eksamen
- De obligatoriske opgaver ligner eksamen og må besvares i grupper op til 3 personer

- 12-timers tag hjem eksamen
- Gruppeeksamen i grupper på max 3 studerende
- Tag-hjem eksamen
  - Tager udgangspunkt i en specifik problemstilling (ofte baseret på en artikel)
  - Anvender simulerede, men realitiske data
  - Opgaven er en blanding af specifikke og åbne spørgsmål
  - Ofte er der også en simulationsopgave
  - Ofte er der en "bevis opgave"
- Besvarelsen skal have et format som en rapport (max 8 sider)
- Sidste år med 12-timers eksamen. Næste år mundtlig eksamen eller 4 timers closed book eksamen.
- Link til gamle eksamensopgaver på Absalon.

## Økonomisk teori og økonometriske metoder

- Opstilling af model
- Kausale sammenhænge
- Forskningsdesign
- Vurdering af centrale antagelser
- Fortolkning af parametre

## Estimatorenes egenskaber

- Beviser og udledninger
- Simulationsopgaver

## Data og **Python**

- Implementering af økonometrisk model
- Estimation
- Hypoteseprøvning