



This Photo by Unknown Author is licensed under [CC BY-SA-NC](#)

F5. SOK-2011: Økonomisk vekst

Empirisk analyse – Hva og hvorfor?

Dette er ikke et statistikk-kurs.

Forelesningen har til mål å gi:

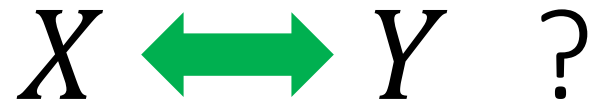
- En grov oversikt/innføring til empirisk analyse
- Støtte til innleveringsoppgaven på kurset

Dere finner mer støtte i appendiks C til kapittel 8 i pensumboken

Hvorfor skal vi bruke empirisk analyse?

Enkelt (og generelt) svar:

Vi ønsker å finne ut om det finns en sammenheng mellom to type observasjoner



Eksempler:

Er det et sammenheng mellom...

- ★ Salg av is og antallet drukningsulykker?
- ★ Eksport av norsk laks og reklamekampanjer for norsk laks rettet mot utlandet?
- ★ Økonomisk aktivitet og mengde CO₂ i atmosfæren?
- ★ Humankapital og økonomisk vekst?

Hvorfor skal vi bruke empirisk analyse?

Til vanlig ønsker vi ikke *bare* å finne ut om det er en korrelasjon mellom to type observasjoner.

Vi har en teori om at en variabel påvirker, eller kan hjelpe til å «forklare» variasjonen i en annen variabel. Med andre ord har vi en **hypotese**, som vi ønsker å **teste**.



Eksempler:

Er det slik at...

- ★ Det å spise is **fører til en økt sannsynlighet** for å drukne?
- ★ Reklamekampanjer rettet mot utlandet **fører til økt** eksport av norsk laks?
- ★ Økonomisk aktivitet **fører til økt** mengde CO2 i atmosfæren?
- ★ Økt humankapital **fører til høyere** vekst i BNP per innbygger?

5 steg for å teste en økonomisk teori

1. Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
2. Operasjonalisere endogene og eksogene variabler, og finn data
3. Velg statistisk metode og analysere data
4. Tolke resultatene
5. Diskutere svakheter i modellen og policy implikasjoner fra resultatene

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

Eksempel 2:

- ★ Solow-modellen uten teknologisk utvikling predikerer at spare-raten påvirker nivået på produksjon per innbygger positivt, og at vekstraten i arbeidskraften påvirker nivået på produksjon per innbygger negativt.
- ★ **Hypotese 1:** Land med høy sparerate vil ha høyere produksjon per innbygger enn land med lav sparerate
- ★ **Hypotese 2:** Land med høy vekstrate i arbeidskraften vil ha lavere produksjon per innbygger enn land med lav vekstrate i arbeidskraften

Endogen variabel:

Nivå på BNP per innbygger

Eksogen variabler:

Sparerate, vekstrate i arbeidskraften

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

Eksempel 3:

- ★ Konvergens-teorien uten teknologisk utvikling predikerer at nivået på BNP per innbygger vil konvergere på lang sikt.
- ★ **Hypotese:** Land som var relativt fattige i starten av «tidsperioden» vil vokse raskere enn land som var rike ved dette tidspunkt.

Endogen variabel:

Gjennomsnittlig årlig vekst i BNP per innbygger

Eksogen variabler:

Nivå på BNP per innbygger i år «null», sparerate, vekstrate i arbeidskraften

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

SOK-2011:

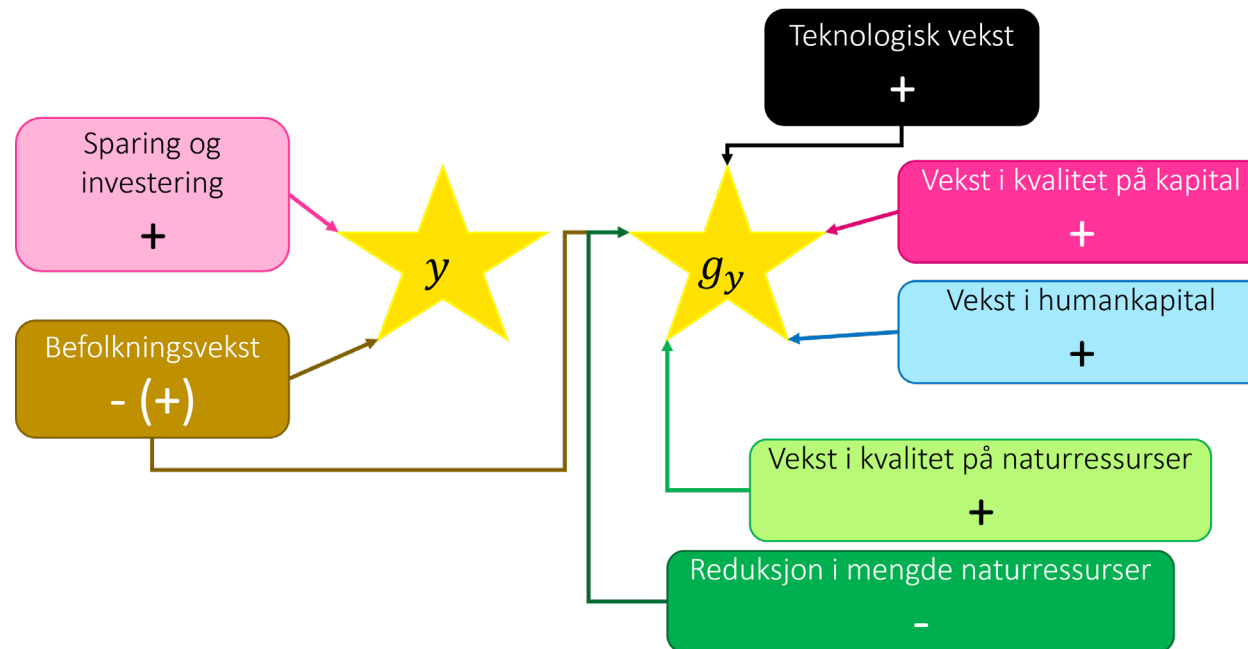
- ★ Hva predikerer Solow-modellen med teknologi og naturressurser?
- ★ Hvilke hypoteser ønsker vi å teste, basert på Solow-modellen?
- ★ Hvilke er våre endogene og eksogene variabler?

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler



5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

Beskriv sambandet så detaljert som mulig.

★ Hvordan henger de eksogene variablene sammen med de endogene variablene?

➡ Ta fram en ligning som beskriver din(e) hypotese(r)

★ Er det andre ting som vi må ta hensyn til, for å unngå å dra feil konklusjoner?

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler



Ta fram en ligning som beskriver din(e) hypotese(r)

Hvordan skal vi beskrive sambandet for Solow-modellen?

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler



Ta fram en ligning som beskriver din(e) hypotese(r) for Solow-modellen

$$\star g_y^{ss} = \frac{1}{1-\alpha} \cdot (g_A + \alpha j + \beta m + \gamma h) - \frac{\gamma}{1-\alpha} (u + n)$$

$$\star g_y = (g_A + \alpha j + \beta m + \gamma h) + \alpha \left(\frac{sy(t) - n \cdot k}{k(t)} \right) - \gamma(u + n)$$

Problem 1

Dette er vekstraten i BNP per **arbeider**, ikke per innbygger!

5 steg for å teste en økonomisk teori

Vekst i BNP per innbygger vs vekst i BNP per arbeider

$$y^L(t) = \frac{Y(t)}{L(t)} = \frac{Y(t)}{P(t)} \cdot \frac{P(t)}{L(t)} = y^P(t) \cdot \frac{P(t)}{L(t)} \quad \Rightarrow \quad g_y^L = g_y^P + \underbrace{\frac{\frac{\partial P(t)}{\partial t}}{P(t)}}_p - n$$

$$g_y^L = g_y^P + p - n \quad \Rightarrow \quad g_y^P = g_y^L - p + n$$

5 steg for å teste en økonomisk teori

Vekst i BNP per innbygger i steady state

$$(g_y^P)^{ss} = \frac{1}{1-\alpha} (g_A + \alpha j + \beta m + \gamma h) - \frac{\gamma}{1-\alpha} (u + n) + n - p$$

$\underbrace{\hspace{15em}}_{(g_y^L)^{ss}}$

Vi kan skrive om dette utrykke slik at vi ikke har n på flere plasser

$$(g_y^P)^{ss} = \frac{1}{1-\alpha} (g_A + \alpha j + \beta m + \gamma h) - \frac{\gamma}{1-\alpha} u + \left(1 - \frac{\gamma}{1-\alpha}\right) n - p$$

5 steg for å teste en økonomisk teori

Skriv om slik at vi bare har n på én plass (slik at vi kan estimere en koeffisient)

Vekst i BNP per innbygger i steady state

$$(g_y^P)^{ss} = \frac{1}{1-\alpha} (g_A + \alpha j + \beta m + \gamma h) - \frac{\gamma}{1-\alpha} u + \left(1 - \frac{\gamma}{1-\alpha}\right) n - p$$

NB:

$$1 - \frac{\gamma}{1-\alpha} = \frac{1-\alpha}{1-\alpha} - \frac{\gamma}{1-\alpha} = \frac{1-\alpha-\gamma}{1-\alpha}$$

$$\alpha + \beta + \gamma = 1 \rightarrow 1 - \alpha - \gamma = \beta$$

$$(g_y^P)^{ss} = \frac{1}{1-\alpha} (g_A + \alpha j + \beta m + \gamma h) - \frac{\gamma}{1-\alpha} u + \frac{\beta}{1-\alpha} n - p$$

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

Empirisk prediksjon fra Solow-modellen i steady state

$$(g_y^P)^{ss} = \frac{1}{1-\alpha} (g_A + \alpha j + \beta m + \gamma h) - \frac{\gamma}{1-\alpha} u + \frac{\beta}{1-\alpha} n - p$$

Vekstraten i BNP per innbygger påvirkes **positivt** av:

- Vekstraten i teknologien: g_A
- Vekstraten i kvaliteten til kapitalen: j
- Vekstraten i kvaliteten til arbeid: m
- Vekstraten i kvaliteten til naturressurser: h
- **Vekstraten i arbeidskraften (!!): n**

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

Empirisk prediksjon fra Solow-modellen i steady state

$$(g_y^P)^{ss} = \frac{1}{1-\alpha} (g_A + \alpha j + \beta m + \gamma h) - \frac{\gamma}{1-\alpha} u + \frac{\beta}{1-\alpha} n - p$$

Vekstraten i BNP per innbygger påvirkes **negativt** av:

- Reduksjonsrate in naturressurser: u
- Vekstraten i befolkningen: p

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

Empirisk spesifikasjon til Solow-modellen i steady state

$$(g_y^P)_i^{ss} = \delta_1 g_{A,i} + \delta_2 j_i + \delta_3 m_i + \delta_4 h_i + \delta_5 u_i + \delta_3 n_i + \delta_6 p_i + \epsilon_i$$

Der: $\delta_1 = \frac{1}{1-\alpha}$, $\delta_2 = \frac{\alpha}{1-\alpha}$, $\delta_3 = \frac{\beta}{1-\alpha}$, $\delta_4 = \frac{\gamma}{1-\alpha}$, $\delta_4 = -\frac{\gamma}{1-\alpha}$, $\delta_5 = -1$

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

Empirisk spesifikasjon til Solow-modellen utenom steady state

$$g_y^P = \underbrace{(g_A + \alpha j + \beta m + \gamma h) + \alpha \left(\frac{sy(t) - n \cdot k}{k(t)} \right) - \gamma(u + n) + n - p}_{g_y^L}$$

Samme her: Vi ønsker å kun ha n på én plass (slik at vi kan estimere en koeffisient).

5 steg for å teste en økonomisk teori

Skriv om slik at vi bare har n på én plass (slik at vi kan estimere en koeffisient)

$$g_y^P = (g_A + \alpha j + \beta m + \gamma h) + \alpha \left(\frac{sy(t) - n \cdot k}{k(t)} \right) - \gamma(u + n) + n - p$$

Steg 1:

Konstant skalautbytte: $\alpha + \beta + \gamma = 1 \rightarrow \gamma = 1 - \alpha - \beta$

➔
$$g_y^P = (g_A + \alpha j + \beta m + \gamma h) + \alpha \left(\frac{sy(t) - n \cdot k}{k(t)} \right) - \gamma u - (1 - \alpha - \beta) \cdot n + n - p$$

5 steg for å teste en økonomisk teori

Skriv om slik at vi bare har n på én plass (slik at vi kan estimere en koeffisient)

$$g_y^P = (g_A + \alpha j + \beta m + \gamma h) + \alpha \left(\frac{sy(t) - n \cdot k}{k(t)} \right) - \gamma u - (1 - \alpha - \beta) \cdot n + n - p$$

Steg 2:

$$\alpha \left(\frac{sy(t) - n \cdot k}{k(t)} \right) = \alpha \cdot s \cdot \frac{y(t)}{k(t)} - \alpha \cdot n$$

➔
$$g_y^P = (g_A + \alpha j + \beta m + \gamma h) + \alpha s \frac{y(t)}{k(t)} - \alpha n - \gamma u - (1 - \alpha - \beta) \cdot n + n - p$$

5 steg for å teste en økonomisk teori

Skriv om slik at vi bare har n på én plass (slik at vi kan estimere en koeffisient)

$$g_y^P = (g_A + \alpha j + \beta m + \gamma h) + \alpha s \frac{y(t)}{k(t)} - \alpha n - \gamma u - (1 - \alpha - \beta) \cdot n + n - p$$

Steg 3:

$$-\alpha \cdot n - (1 - \alpha - \beta) \cdot n + n = (-\alpha - 1 + \alpha + \beta + 1) \cdot n = \beta \cdot n$$

➔
$$g_y^P = (g_A + \alpha j + \beta m + \gamma h) + \alpha s \frac{y(t)}{k(t)} - \gamma u + \beta \cdot n - p$$

5 steg for å teste en økonomisk teori

$$g_y^P = (g_A + \alpha j + \beta m + \gamma h) + \alpha s \frac{y(t)}{k(t)} - \gamma u + \beta \cdot n - p$$

Problem 2:

$$\rightarrow y(t) = A_0 \cdot e^{(g_A + \alpha j + \beta m + \gamma h) \cdot t} K(t)^\alpha \cdot L(t)^{\beta-1} \cdot R_0^\gamma \cdot e^{-\gamma u t}$$

$$y(t) = f(A_0, g_A, L_0, n, R_0, j, m, h, u, \alpha, \beta, \gamma, s, t)$$

$$\rightarrow k(t) = v(A_0, g_A, L_0, n, R_0, j, m, h, u, \alpha, \beta, \gamma, s, t)$$

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

Empirisk spesifikasjon til Solow-modellen utenom steady state

$$g_y^P = (g_A + \alpha j + \beta m + \gamma h) + \alpha s \frac{y(t)}{k(t)} - \gamma u + \beta n - p$$

Vekstraten i i BNP per innbygger påvirkes **positivt** av:

- Vekstraten i teknologien: g_A
- Vekstraten i kvaliteten til kapitalen: j
- Vekstraten i kvaliteten til arbeid: m
- Vekstraten i kvaliteten til naturressurser: h
- spareraten: s
- Vekstraten i arbeidskraften: n
- Nivået på $\frac{y(t)}{k(t)}$

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

Empirisk spesifikasjon til Solow-modellen utenom steady state

$$g_y^P = (g_A + \alpha j + \beta m + \gamma h) + \alpha s \frac{y(t)}{k(t)} - \gamma u + \beta n - p$$

Vekstraten i BNP per innbygger påvirkes **negativt** av:

- Reduksjonsrate in naturressurser: u
- Befolkningsveksten: p

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

Empirisk spesifikasjon til Solow-modellen utenom steady state

$$g_{y_i}^P = \delta_1 g_{A,i} + \delta_2 j_i + \delta_3 m_i + \delta_4 h_i + \delta_5 u_i + \delta_3 n_i + \delta_6 p_i + \delta_7 \cdot s \frac{y(t)}{k(t)} + \epsilon_i$$

5 steg for å teste en økonomisk teori

1.

Bruk en økonomisk teori for spesifisere din(e) hypotese(r)
(matematisk, eller ved bruk av økonomiske argumenter)

Beskriv et hypotetisk samband mellom eksogene (forklarende) og endogene (avhengige) variabler

Empiriske spesifikasjoner til Solow-modellen i og utenom steady state

$$(g_y^P)_i^{ss} = \delta_1 g_{A,i} + \delta_2 j_i + \delta_3 m_i + \delta_4 h_i + \delta_5 u_i + \delta_3 n_i + \delta_6 p_i + \epsilon_i$$

$$g_y^P = \delta_1 g_{A,i} + \delta_2 j_i + \delta_3 m_i + \delta_4 h_i + \delta_5 u_i + \delta_3 n_i + \delta_6 p_i + \delta_7 \cdot s \cdot \frac{y(t)}{k(t)} + \epsilon_i$$

5 steg for å teste en økonomisk teori

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

Fire spørsmål:

1. Hvor kan vi finne data til modellen?
2. Hvordan skal vi håndtere $\frac{y(t)}{k(t)}$?
3. Hvordan kan vi måle $g_{A,i}$, j_i , m_i , h_i , u ?
4. Har vi glemt noe som åpenbart påvirker vekstraten i produksjon per innbygger, men som vi ikke har med?

5 steg for å teste en økonomisk teori

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

1. Hvor kan vi finne data til modellen?

Målbilde:

	g_y	g_A	j	m	h	u	n	p	$s \cdot \frac{y(t)}{k(t)}$
Land 1 (f.eks. Norge)	0.015	0.018	0.05	0.001	0.05	0.003	0.005	0.002	75
Land 2 (f.eks. USA)	0.013	0.02	0.02	0.002	0.02	0.006	0.008	0.002	89
Land 3 (f.eks. Malawi)	0.052	0.008	0.04	0.03	-0.01	0.002	0.04	0.05	30
Land 4 (f.eks. Thailand)	0.03	0.04	0.05	0.035	-0.02	0.002	0.05	0.06	40
Land 5 (f.eks. Zimbabwe)	-0.001	0.001	0.002	0.0001	0.001	0.001	0.03	0.04	20

5 steg for å teste en økonomisk teori

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

1. Hvor kan vi finne data til modellen?

Det finns mange datakilder med informasjon om variabler knyttet til nivå på og vekst i BNP per innbygger

Eksempler:

- ★ Penn-World tables (PWT)
- ★ World Development Indicators (WDI)

Både PWT og WDI finns tilgjengelige i R-studio som pakker

Pensumboken bruker WDI

5 steg for å teste en økonomisk teori

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

1. Hvor kan vi finne data til modellen?

★ Word Development Indicators (WDI): <https://datatopics.worldbank.org/world-development-indicators/>

- Et svært dataset med variabler knyttet til økonomisk utvikling.
- Satt sammen av Verdensbanken

- I Rstudio heter pakken «wdi»
- Søkbar:
 - `WDIsearch(string = "name", field = "name", short = TRUE, cache = NULL)`

F.eks. "Net saving"

Vi vil bruke seminaret til å bli kjent med datasetet

5 steg for å teste en økonomisk teori

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

1. Hvor kan vi finne data til modellen?

WDI inneholder data for hvert land på årlig basis



Vi ønsker å se på de «fundamentale bestemmelsesfaktorene» til økonomisk vekst, ikke svært kortsiktige svingninger.



Bruk gjennomsnittlige verdier for tidsperioden (2000-2019)

5 steg for å teste en økonomisk teori

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

2. Hvordan skal vi håndtere $s \cdot \frac{y}{k}$?

★ Dersom vi finner data på $s, Y(t), K(t)$ og $L(t)$ kan vi egentlig beregne dette.

★ MEN! Dersom g_y øker, vil dette påvirke både $y(t)$ og $k(t)$. Vi lager et endogenitetsproblem!

«Løsning»:

➔ Da s er den sentrale parameteren her, foreslår pensumboken at vi fokuserer på denne.

5 steg for å teste en økonomisk teori

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

3. Klarer vi å måle g_A, j, m, h, u ?

★ u

WDI inneholder to variabler knyttet til bruk av naturressurser:

1. NY.ADJ.DRES.GN.ZS - Adjusted savings: natural resources depletion (% of GNI)
2. NY.GDP.TOTL.RT.ZS - Total natural resources rents (% of GDP)

★ m

Finns dessverre ingen instrumenter som måler dette (så hvitt jeg vet). Vi kan bruke en «proxy»: Gjennomsnittlig antall år i skole i befolkningen.

BAR.SCHL.15UP -Barro-Lee: Average years of total schooling, age 15+, total

Kan også prøve å beregne vekst i kvalitet i skole

SPI.D3.4.EDUC -GOAL 4: Quality Education (5 year moving average)

5 steg for å teste en økonomisk teori

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

3. Klarer vi å måle g_A, j, m, h, u ?



g_A, j, h

Finns dessverre ingen instrumenter som måler dette (så hvitt jeg vet).

Pragmatisk løsning:

$$(g_y^P)_i^{SS} = \alpha + \delta_3 \text{skoleår}_i + \delta_5 u_i + \delta_3 n_i + \delta_6 p_i + \delta_7 \cdot s_i + \epsilon_i$$

Teknologi og annet som vi ikke observerer men som er likt for alle land blir fanget opp her

Forskjeller i teknologi mellom ulike land blir fanget opp her

5 steg for å teste en økonomisk teori

Konvergens!!!

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

4. Har vi glemt noe som åpenbart påvirker vekstraten i produksjon per innbygger, men som vi ikke har med?



Solow-modellen er basert på «lang sikt»

Alle markeder er i likevekt

Både nivå på produksjonen og vekst i produksjonen drivs av produksjonsmulighetene (tilbudet)



Solow-modellen antar at økonomien er lukket

På kort og midlertidlig sikt vil etterspørselen påvirke vekstraten i y .

På kort og midlertidlig sikt vil ikke alle markeder være i likevekt

$$S \neq I$$

De fleste økonomier handler med andre land
Både investeringer fra utlandet og handel med utlandet gir tilgang til ny teknologi

5 steg for å teste en økonomisk teori

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

4. Har vi glemt noe som åpenbart påvirker vekstraten i produksjon per innbygger, men som vi ikke har med?

★ Konvergensteorien predikerer at fattige land «komme i kapp» rike land.

➔ Fattigere land vil vokse raskere enn rike land

Hvordan kan vi teste dette?

5 steg for å teste en økonomisk teori

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

4. Har vi glemt noe som åpenbart påvirker vekstraten i produksjon per innbygger, men som vi ikke har med?

Konklusjon: Kan vurdere å ta med variabler som måler vekst i handel og vekst i investeringer, og opprinnelig nivå på BNP per innbygger.

★ g_I

WDI inneholder en variabel som måler vekst i investeringer

NE.GDI.FTOT.KD.ZG - Gross fixed capital formation (annual % growth)

★ g_X

WDI inneholder en variabel som måler vekst i eksport

NE.EXP.GNFS.KD.ZG - Exports of goods and services (annual % growth)

★ $BNP_{pc,2000}$

WDI inneholder en variabel som måler nivå på BNP per innbygger i PPP-dollar

NY.GDP.PCAP.PP.KD - GDP per capita, PPP (constant 2017 international \$)

5 steg for å teste en økonomisk teori

2.

Operasjonalisering

Hvilken informasjon kan vi bruke for å representere variablene i vår(e) hypotese(r)?

“Korrekt” spesifikasjon av modellen

$$g_{y_i}^P = \delta_1 g_{A,i} + \delta_2 j_i + \delta_3 m_i + \delta_4 h_i + \delta_5 u_i + \delta_6 n_i + \delta_7 p_i + \delta_7 \cdot s \cdot \frac{y(t)}{k(t)} + \epsilon_i$$

Pragmatisk spesifikasjon av modellen

$$g_{y_i}^P = \alpha + \vartheta_1 \text{Skoleår}_i + \vartheta_2 \text{NRY}_i + \vartheta_3 n_i + \vartheta_4 p_i + \vartheta_5 \text{NSY}_i + \vartheta_6 g_{I_i} + \vartheta_7 g_{X_i} + \vartheta_8 \text{BNP}_{2000_i} + \epsilon_i$$

NB: Dersom dere finner variabler som dere mener burde være med i modellen, og kan gi økonomiske argumenter til å inkludere dem, så kan dere inkludere disse i modellen.

5 steg for å teste en økonomisk teori

3.

Velg statistisk metode og analysere data

Hvilken metode skal vi bruke til å estimere modellen?

Valg av estimerings-metode avhenger både egenskapene til den avhengige variabelen, hvordan de forklarende variablene kan tenkes henge sammen med den avhengige variabelen og andre forklarende variabler.

Den enkleste typen av empirisk model: **Ordinary Least Squares** (minste kvadratt-metoden)

5 steg for å teste en økonomisk teori

3.

Velg statistisk metode og analysere data

Hvilken metode skal vi bruke til å estimere modellen?

Ordinary Least Squares (minste kvadratt-metoden)

Antakelser:

- ★ Den avhengige variabelen er kontinuerlig, kardinal og har en normalfordeling
- ★ Funksjonen er linjer i parameterne $y_i = \alpha + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \beta_3 X_{2,i}^2 + \epsilon_i$
- ★ Utvalget er tilfeldig (variablene er i.i.d – uavhengig og identisk fordelte)
- ★ De forklarende variablene er ikke (perfekt) korrelerte med hverandre
- ★ Feiltermene er homoskedastiske
- ★ Det er usannsynlig at det er ekstreme observasjoner (outliers) i utvalget

5 steg for å teste en økonomisk teori

3.

Velg statistisk metode og analysere data
Hvilken metode skal vi bruke til å estimere modellen?

Eksempler: Endogene variabler med ulike egenskaper

★ Vekstraten i BNP per innbygger	Kontinuerlig, kardinal (2% er dobbelt så mye som 1%), kan ta både positive og negative verdier.
★ Lønn	Kontinuerlig, kardinal (500 NOK er dobbelt så mye som 250 NOK), observasjoner kan bare ta positive verdier. Vi observerer ikke folk som ikke har jobb (seleksjon).
★ Fornøydhet med livet (0-10)	Ordinal skala (10 er ikke sikkert dobbelt så mye som 5)
★ Skilsmisse	Dikotom variabel (1/0), ordinal, ønsker å få fram sannsynlighet (kan ikke være høyere enn 100% eller lavere enn 0%)

Funker fint med OLS

Funker mindre bra med OLS

5 steg for å teste en økonomisk teori

3.

Velg statistisk metode og analysere data

Hvilken metode skal vi bruke til å estimere modellen?

Steg 1: Se på data (deskriptiv analyse)

★ Ta fram verdier for gjennomsnitt (eller median), minste verdier, største verdier og spredning (standardavvik)

➔ Dette gir oss et første bilde av datamaterialet. Vi kan se om det er noen verdier som ser «rare» ut

Variabel	N	Gjennomsnitt	Standard avvik	Min	Maks
BNP per innbygger (PPP)	156	230521.123	21304.175	1050.222	125365.253
Vekstrate BNP per innbygger (%)	156	2.326	4.231	-0.123	14.085
Nettosparing (% av BNI)	72	9.158	14.189	4.298	34.596
Vekst i investeringer (%)	156	5.596	11.156	-10.963	131.312
Befolkningsvekst (%)	156	2.963	1.723	0.001	5.597

5 steg for å teste en økonomisk teori

3.

Velg statistisk metode og analysere data
Hvilken metode skal vi bruke til å estimere modellen?

Steg 2: Velg estimerings-metode og estimere modellen

På dette kurset skal dere bruke OLS (minste kvadratmetoden).

5 steg for å teste en økonomisk teori

3.

Velg statistisk metode og analysere data
Hvilken metode skal vi bruke til å estimere modellen?

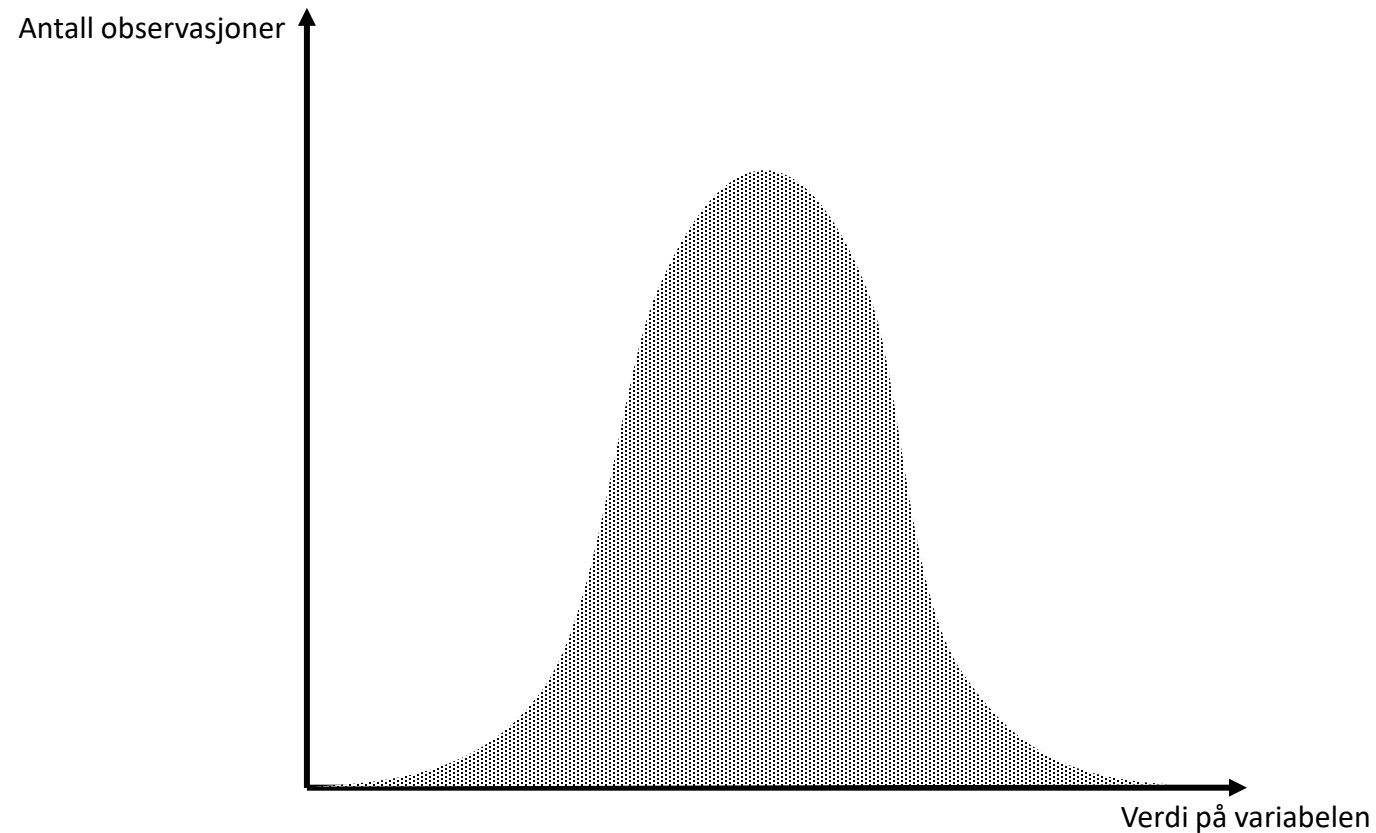
Ordinary Least Squares (minste kvadratt-metoden)

Antakelser:

- ★ Den avhengige variabelen er kontinuerlig, kardinal og har en normalfordeling
- ★ Funksjonen er linjer i parameterne $y_i = \alpha + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \beta_3 X_{2,i}^2 + \epsilon_i$
- ★ Utvalget er tilfeldig (variablene er i.i.d – uavhengig og identisk fordelte)
- ★ De forklarende variablene er ikke (perfekt) korrelerte med hverandre
- ★ Feiltermene er homoskedastiske
- ★ Det er usannsynlig at det er ekstreme observasjoner (outliers) i utvalget

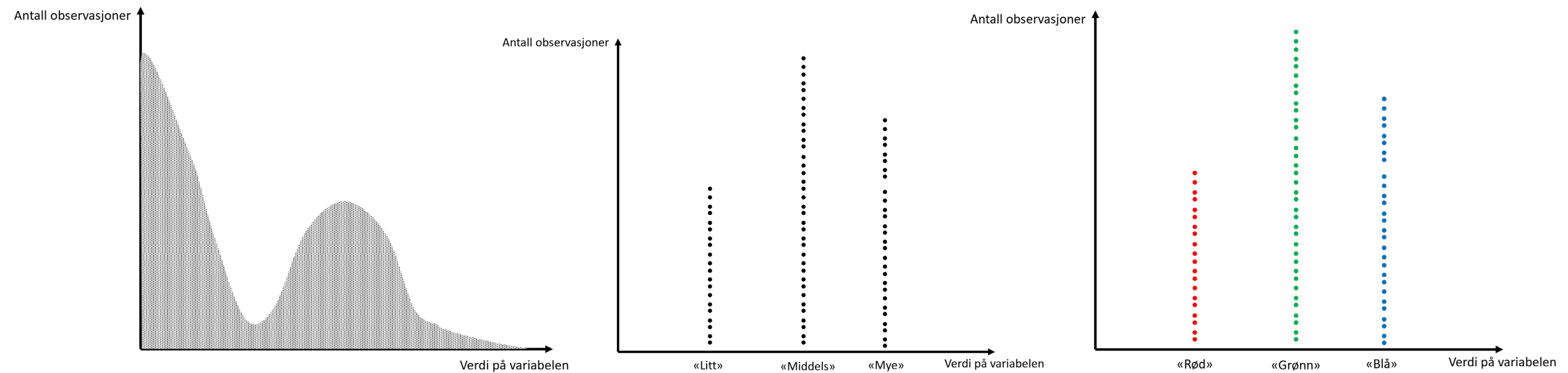
Ordinary Least Squares

- ★ Den avhengige variabelen er kontinuerlig, kardinal og har en normalfordeling



Ordinary Least Squares

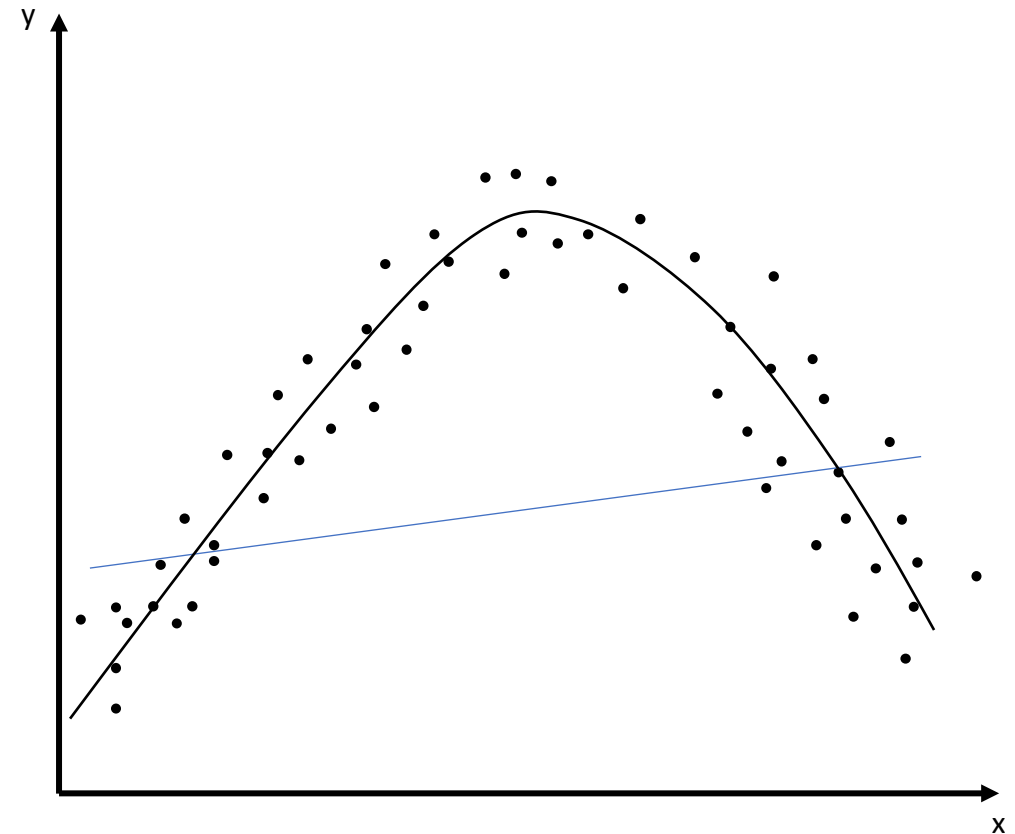
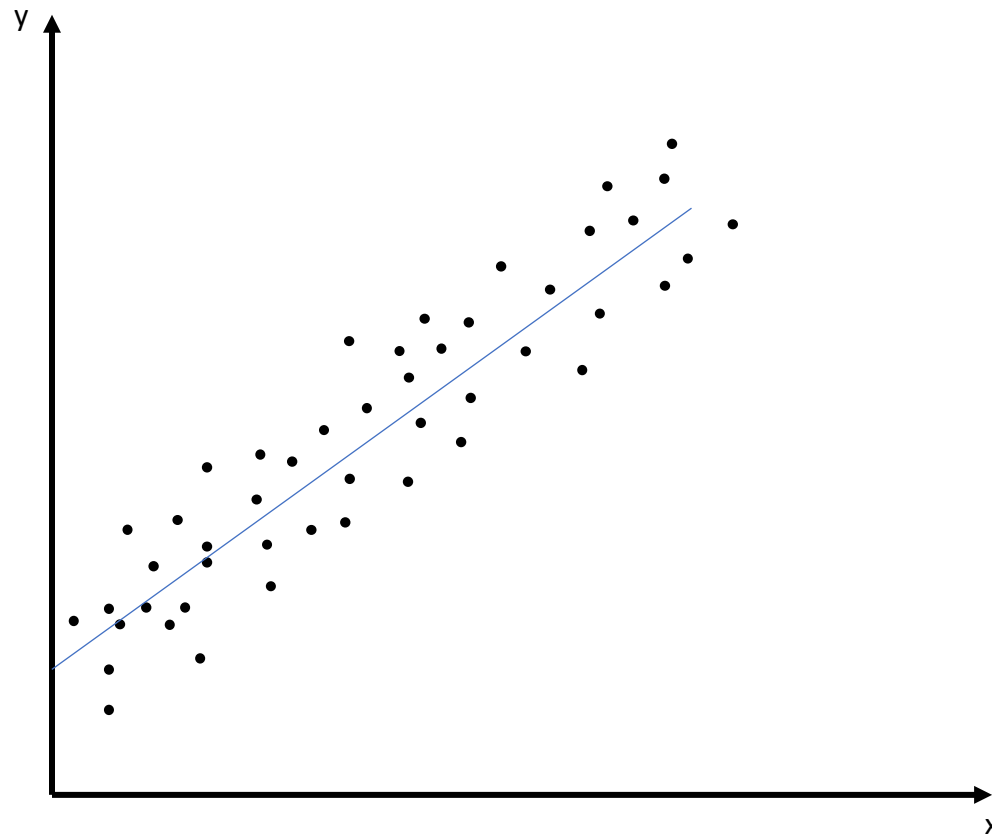
- ★ Den avhengige variabelen er IKKE kontinuerlig, kardinal eller har en normalfordeling



Ordinary Least Squares

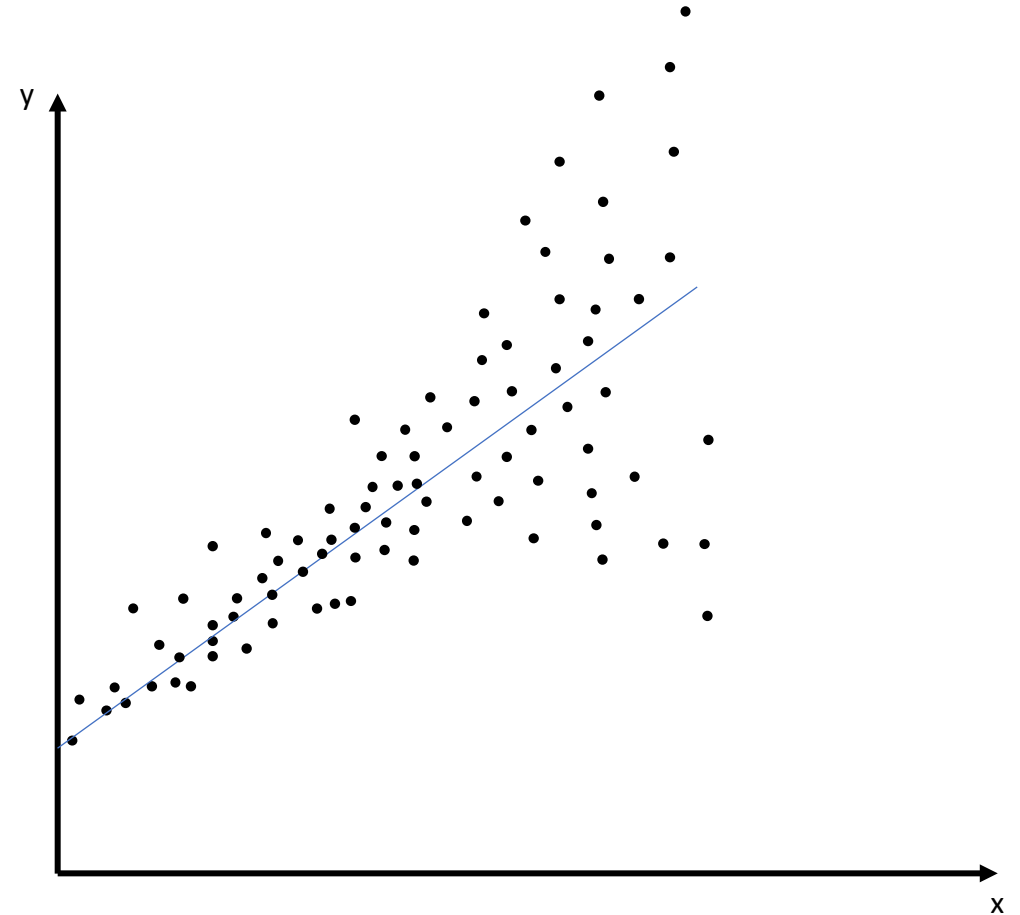
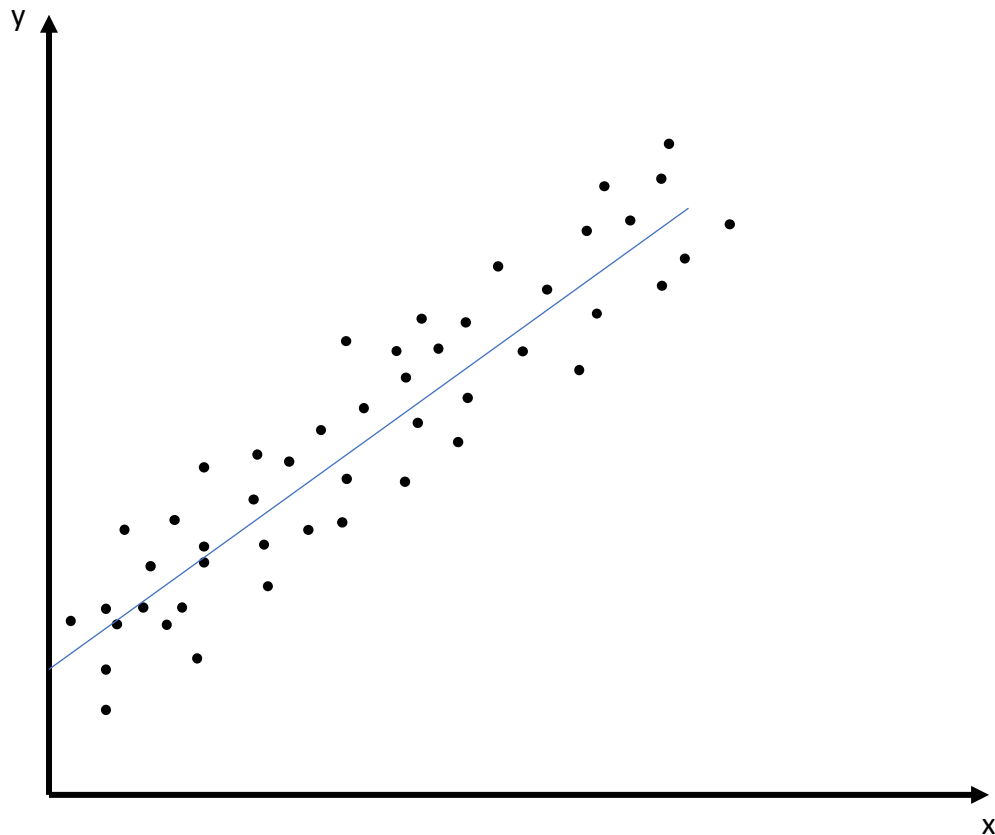
★ Den avhengige variabelen er kontinuertlig, kardinal og har en normalfordeling

★ Funksjonen er linjer i parameterne $y_i = \alpha + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \beta_3 X_{2,i}^2 + \epsilon_i$



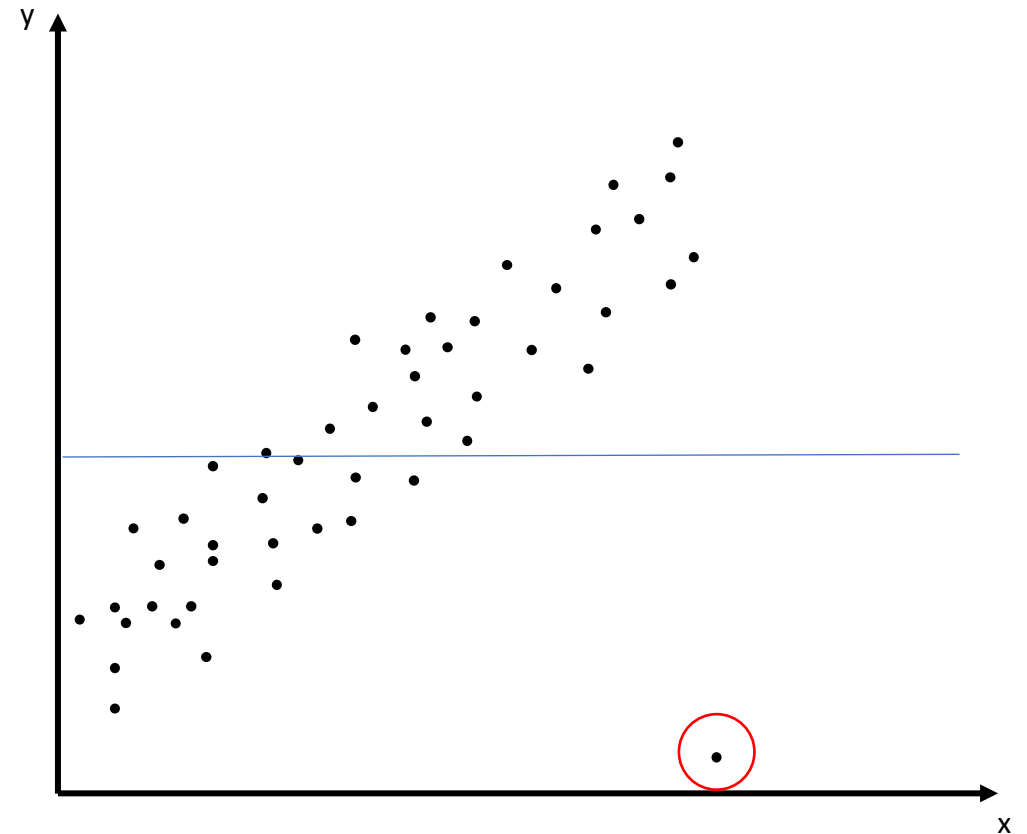
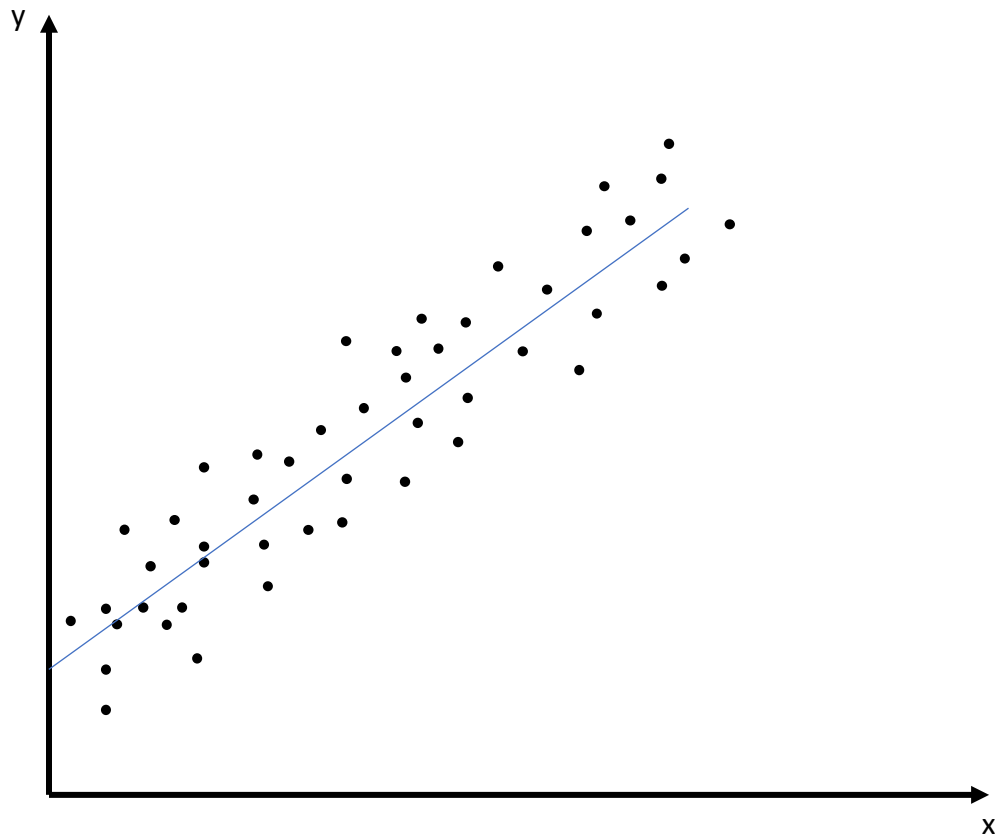
Ordinary Least Squares

★ Feiltermene er homoskedastiske



Ordinary Least Squares

- ★ Det er usannsynlig at det er ekstreme observasjoner (outliers) i utvalget



Ordinary Least Squares

- ★ De forklarende variablene er ikke (perfekt) korrelerte med hverandre

$$BNP_{2000} = \alpha + \beta_1 ForventetLevealder_{2000} + \beta_2 HIVPrevalens_{2000} + \beta_3 Sparing_{2015} + \dots + \epsilon$$

5 steg for å teste en økonomisk teori

4.

Tolke resultatene

Hva betyr de resultatene som vi har fått fram?

Mock-up eksempel

Variabel	β	Std.error	t value	Pr(> t)
Nettosparing (% av BNI)	0.081	0.0122	6.639	0.001
Vekst i investeringer (%)	0.152	0.037	4.109	0.001
Befolkningsvekst (%)	1.261	4.621	0.273	0.893
Vekst i eksport (%)	0.191	0.089	2.112	0.049
Reduksjon i naturressurser (%)	-0.053	0.0198	-2.714	0.008

N= 103, F-verdi: 21.87 (p =.001), R2 =0.75, adj-R2 = 0.63

5 steg for å teste en økonomisk teori

4.

Tolke resultatene

Hva betyr de resultatene som vi har fått fram?

Variabel	β	Std.error	t value	Pr(> t)
Nettosparing (% av BNI)	0.081	0.012	6.639	0.001
Vekst i investeringer (%)	0.152	0.037	4.109	0.001
Befolkningsvekst (%)	1.261	4.621	0.273	0.893
Vekst i eksport (%)	0.191	0.089	2.112	0.049
Reduksjon i naturressurser (%)	-0.053	0.0198	-2.714	0.008

Beskriv resultatene i ord.

- ★ Hvilke variabler er signifikant korrelert med den avhengige variabelen?
- ★ Er effekten positiv eller negativ (fører en økning i den forklarende variabelen til en økning eller reduksjon i den avhengige variabelen)?
- ★ Går det å si noe om størrelsen på effekten?
F.eks. i henhold til resultatene er en økning i spareraten med 1 prosentpoeng assosiert med en økning i vekstraten i BNP per innbygger med 0.081 prosentpoeng.

5 steg for å teste en økonomisk teori

4.

Tolke resultatene

Hva betyr de resultatene som vi har fått fram?

Variabel	β	Std.error	t value	Pr(> t)
Nettosparing (% av BNI)	0.081	0.012	6.639	0.001
Vekst i investeringer (%)	0.152	0.037	4.109	0.001
Befolkningsvekst (%)	1.261	4.621	0.273	0.893
Vekst i eksport (%)	-0.191	0.089	2.112	0.049
Reduksjon i naturressurser (%)	-0.053	0.0198	-2.714	0.008

Beskriv resultatene i ord.

- ★ Stemmer resultatene med teorien?
- ★ Dersom resultatene ikke stemmer med teorien – er det noen måte å forklare dette på?
 - ★ Korrelasjoner mellom forklaringsvariablene?
 - ★ Ekstreme observasjoner og/eller målefeil?
 - ★ Ikke-observerte faktorer som driver resultatene?

5 steg for å teste en økonomisk teori

4.

Tolke resultatene

Hva betyr de resultatene som vi har fått fram?

Variabel	β	Std.error	t value	Pr(> t)
Nettosparing (% av BNI)	0.081	0.012	6.639	0.001
Vekst i investeringer (%)	0.152	0.037	4.109	0.001
Befolkningsvekst (%)	1.261	4.621	0.273	0.893
Vekst i eksport (%)	-0.191	0.089	2.112	0.049
Reduksjon i naturressurser (%)	-0.053	0.0198	-2.714	0.008



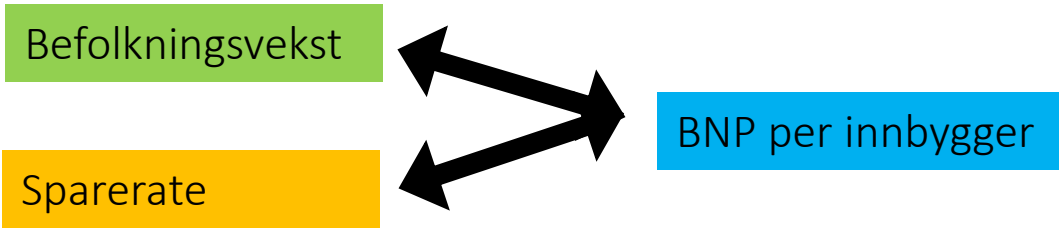
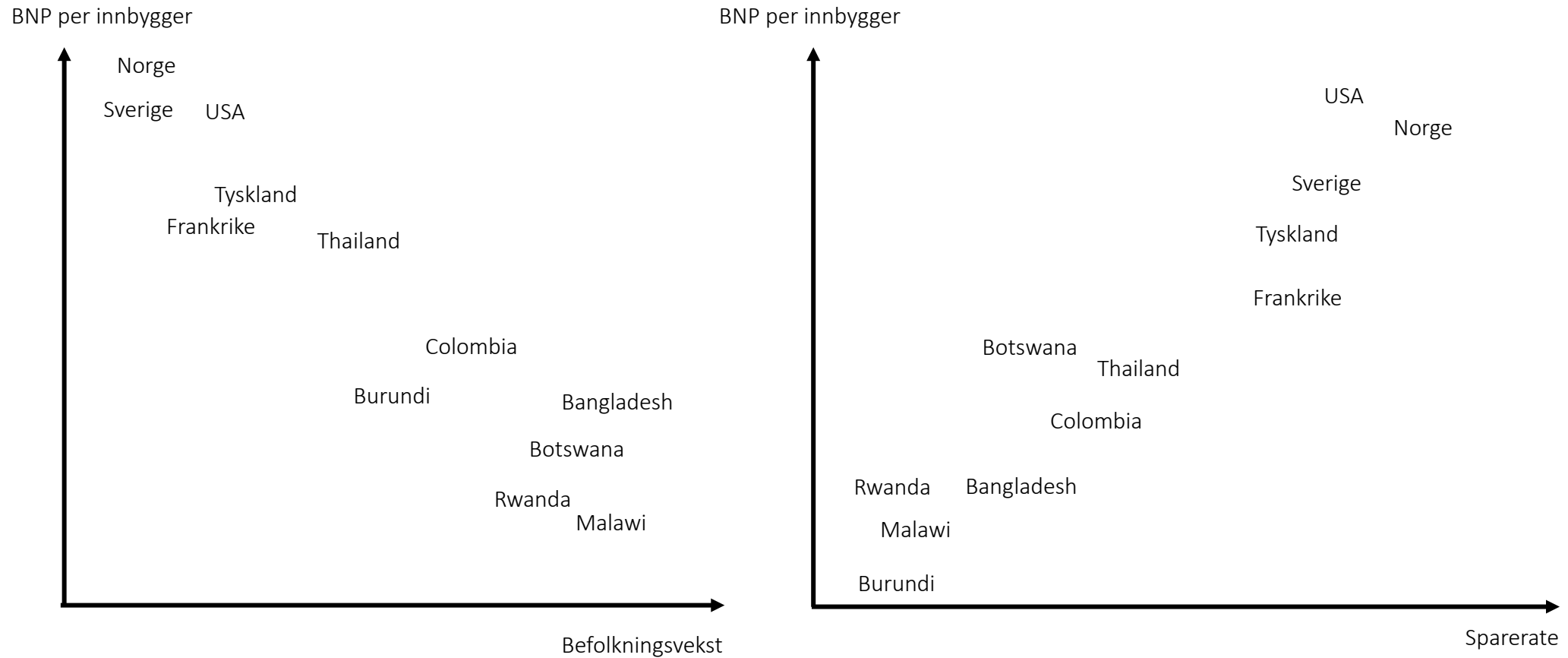
Hvor godt passer modellen til data?



Hvor mye av variansen i den avhengige variabelen klarer de forklarende variablene å forklare?

N= 103, F-verdi: 21.87 (p =.001), R2 =0.75, adj-R2 = 0.63

NB: Innen samfunnsvitenskap kan vi sjelden finne «bevis» for at vår(e) hypotese(r) er «sann(e)»!

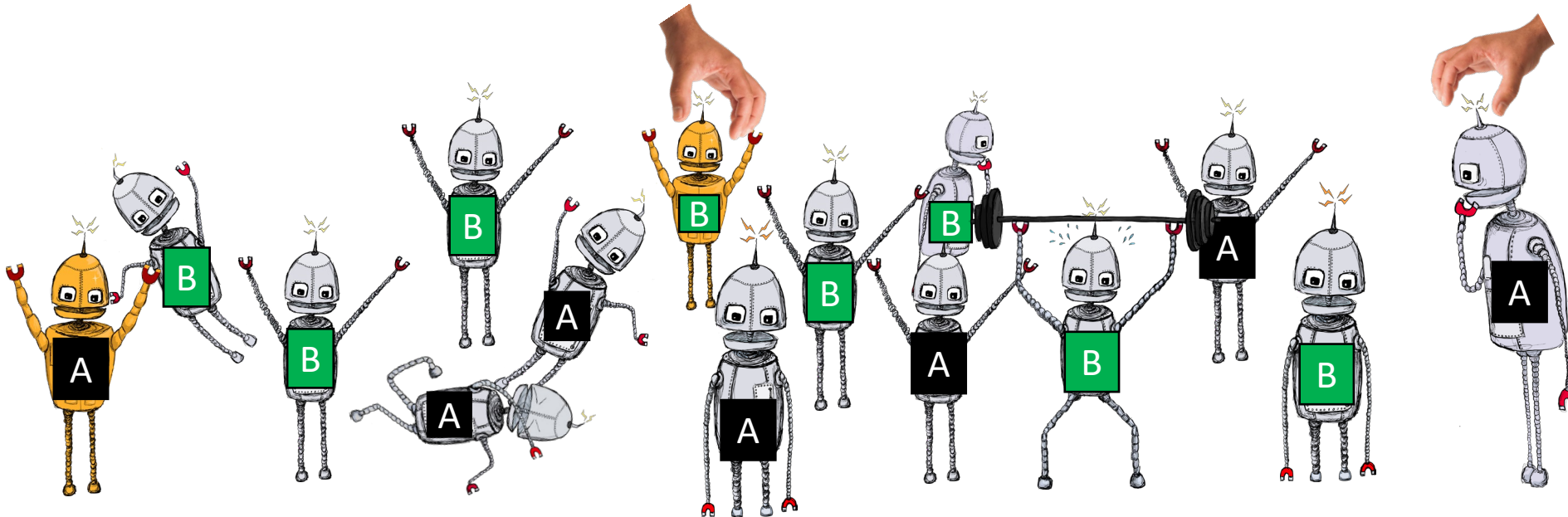


NB: Innen samfunnsvitenskap kan vi sjelden finne «bevis» for at vår(e) hypotese(r) er «sann(e)»!

Unntak:

Studier i mikroøkonomi: I noen situasjoner kan vi lage randomiserte eksperimenter

Studier i makroøkonomi: I noen situasjoner kan vi benytte «naturlige» eksperimenter



Vi har ikke tilgang til naturlige eksperimenter → våre resultater viser korrelasjoner, ikke årsaks-sammenhenger

5 steg for å teste en økonomisk teori

5.

Diskutere svakheter i modellen og policy implikasjoner fra resultatene



Hvilke svakheter finns det i den empiriske analysen?

F.eks. Mangler vi data på viktige variabler?

Kan det være problemer med den valgte estimeringsmetoden?



Gitt svakhetene i modellen, hva skal vi ta med oss fra analysen (hva kan vi og andre lære)?

Dersom noen med makt å påvirke leser din oppgave, hvilken lærdom skal denne ta fra analysen?