



UiT Norges arktiske universitet

Handelshøgskolen ved UiT

Fakultet for biovitenskap, fiskeri og økonomi

Utfordring 2

Empirisk analyse av Solow-modellen

Kandidatnummer: 70, 84, 72

Sok-2011, Vår 2024

Sammendrag

I vår empiriske analyse av Solow-modellen brukte vi data fra World Development Indicators (WDI) for å undersøke forholdet mellom BNP per innbygger og en rekke variabler som spareraten, befolkningsveksten, humankapital, forbruk av naturressurser, og det initielle nivået på BNP per innbygger. Disse variablene ble valgt basert på deres antatte innvirkning på økonomisk vekst i henhold til Solow-modellen.

Spredningsdiagrammene viste svake, men likevel merkbare sammenhenger mellom de uavhengige variablene og BNP per innbygger, i tråd med Solow-modellens prediksjoner. En økning i spareraten indikerte en tendens til økt BNP per innbygger, mens høyere befolkningsvekst syntes å korrelere med lavere BNP per innbygger. Videre illustrerte diagrammene en positiv sammenheng mellom humankapital og BNP per innbygger, samt en negativ effekt av høyt forbruk av naturressurser på økonomisk vekst.

Gjennom regresjonsanalyse bekreftet vi de observerte sammenhengene. Spesielt viste analysen at spareraten og humankapitalen har en positiv effekt på BNP per innbygger, mens befolkningsveksten og forbruket av naturressurser har en negativ effekt. Dette er i samsvar med Solow-modellens forutsigelser om at sparing og investering i humankapital fremmer økonomisk vekst, mens rask befolkningsvekst og overdreven utnyttelse av naturressurser kan hemme denne veksten.

Den videre analysen av vekstraten i BNP per innbygger viste også en interessant konvergensprosess, hvor land med lavere initielt BNP per innbygger opplevde raskere vekst enn de med høyere initielt BNP per innbygger. Dette understreker teorien om økonomisk konvergens og Solow-modellens relevans for å forstå mekanismene bak økonomisk vekst.

Selv om våre resultater stort sett støtter Solow-modellens prediksjoner, må vi erkjenne begrensningene i vår analyse. For det første antar modellene lineære sammenhenger, mens de faktiske forholdene mellom variablene kan være mer komplekse. Videre, selv om vi har observert statistisk signifikante effekter, forklarer modellene kun en del av variansen i BNP per innbygger, noe som tyder på at det er andre viktige faktorer som ikke er inkludert i analysen.

Til tross for disse begrensningene, gir vår analyse verdifull innsikt i de økonomiske mekanismene som påvirker vekst og understreker viktigheten av sparing, investering i humankapital, og bærekraftig forvaltning av naturressurser for å fremme økonomisk utvikling. Disse funnene understreker Solow-modellens relevans og dens bidrag til forståelsen av økonomisk vekst.

Innholdsfortegnelse

Utfordring 2.1	5
Beskriv data-materialet som du vil bruke i analysen.	5
Utfordring 2.2	7
a. Lag spredningsdiagram (scatterplots) som illustrerer sammenhengen mellom spareraten med BNP per innbygger og befolknings-vekstraten med BNP per innbygger.	7
b. Estimer en regresjonsmodell (minste kvadrats-metode) som tester om spar- eraten og befolkningsveksten forklarer variasjon i BNP per innbygger mellom ulike land.	8
c. Tolke resultatene fra spredningsdiagrammen og regresjonsanalysen.	9
Utfordring 2.3	10
a. Spredningsdiagram som viser sammenhengen mellom Humankapital, for- bruket av naturressurser og og initialt nivå på BNP per innbygger mot vekstraten i BNP per innbygger.	10
b. Estimer en regressionsmodell (minste kvadrats-metode) som tester om spar- eraten, befolkningsvekstraten, humankapital, forbruk av naturressurser og initialt nivå på BNP per innbygger, forklarer variasjon i vekstraten i BNP per innbygger i ulike land. Den modellen dere skal estimere kan bli beskrevet av ligningen under (årsaken til at vi bruker logaritmen av y_0 er fremst at BNP per innbygger er mye større enn de andre variablene. Ved å bruke den naturlige logartmen får vi en bedre “fit” av data til modellen:	12
c. Tolke resultatene fra spredningsdiagrammene og regresjonsanalysen, og diskutere eventuelle svakheter eller begrensninger.	13
Referanser	15
Appendix Generell KI bruk	16
Appendix 1.1 KI bruk	18
Appendix Kapittel 1.2 KI bruk	20
Appendix Kapittel 1.3 KI bruk	20

Figurliste

1	Sparerate	7
2	Befolkningsvekst	7
3	Humankapital	10
4	Humankapital logaritmisk BNP per innbygger	10
5	Vekstrate i BNP per innbygger	11
6	Naturressurser	11
7	Naturressurser med log y-akse	11
8	Linearitetstest	14
9	Normalitetstest	14
10	Homoskedasitetstest	14

Tabelliste

1	Spredningsmål	6
2	Regresjonsmodel sparerate og befolkningsvekst påvirkning på bnp	8
3	Multippel lineær regresjonsmodeller med log og ikke log avBNP per Innbygger	12

Kode er tatt fra Mannberg (2023)

Utfordring 2.1

Beskriv data-materialet som du vil bruke i analysen.

I denne utfordringen skal vi beskrive datamaterialet vi vil bruke i vår empiriske analyse i de to neste oppgavene, samt de variablene vi vil bruke. Datamaterialet vi har er data fra WDI som også finnes som en egen pakke i R for at de skal være lett tilgjengelig for analyse. WDI står for World Development Indicators, som er en samling av data som [The World Bank](#) har satt sammen og som ligger åpen for alle å bruke. De har satt sammen disse dataene fra offisielle anerkjente internasjonale kilder og de sørger for at de mest oppdaterte og nøyaktige globale utviklings-dataene er tilgjengelige samt at de inkluderer nasjonale, regionale og globale estimer. Det er denne databasen vi skal bruke, samt at vi skal ta utgangspunkt i den koden vi har fått fra fagansvarlig med de variablene som er satt.

Det er 5 sett med data som vi skal bruke fra WDI.

Det første settet med data er NY.GDP.PCAP.PP.KD som vi endrer navnet på til gdpcc. Ifølge [The World Bank](#) så inneholder dette datasettet en beregnet PPP, som betyr at det er BNP per innbygger som er omregnet til kjøpekraftsparitet. Dette betyr at landene kan direkte sammenlignes på grunn av valuta-justering, slik at etter å ha tatt hensyn til valuta-kurs så får man kjøpt den samme mengden varer i alle land.

Det andre datasettet er HD.HCI.OVRL som vi endrer navn til hci_index, samt hci_fem til HD.HCI.OVRL.FE og hci_male til HD.HCI.OVRL.MA, disse datasettene inneholder beregnet humankapital av hva vi har tolket av innholdet og gitt kode, da vi ikke klarte å finne informasjon på Metadata Glossary. Disse tre blir samlet til et datasett som da har informasjon om både samlet, men også menn og kvinner for seg om hvor god utdanning det er blant disse i de forskjellige landene i en indeks på mellom 0 og 1.

Det tredje datasettet er NY.ADJ.DRES.GN.ZS som vi endrer navn på til nry, dette datasettet inneholder anslaget av samlet bruk av naturressurser slik som gull, olje, kull og natur og deres resterende levetid (25 år) ifølge [The World Bank](#).

Det fjerde datasettet NY.ADJ.NNAT.GN.ZS som vi kaller for savings er ifølge [The World Bank](#) «Net national savings are equal to gross national savings less the value of consumption of fixed capital» som direkte oversatt blir til «Netto nasjonal sparing er lik brutto nasjonal sparing fratrukket verdien av forbruk av fast kapital».

Og det siste datasettet er SP.POP.GROW som vi kaller for popgrowth er data på den eksponentielle vekstraten for midtårsbefolkningen uttrykt i prosent. Denne er også basert på at befolkningen er basert på facto-definisjonen av befolkning som vil si at de teller alle innbyggerne uavhengig av statsborgerskap eller juridisk status ifølge [The World Bank](#).

Disse datasettene blir slått sammen til ett datasett basert på landene og er for tidsperioden 2010 til 2015. Unntak er gdpcc på grunn av manglende observasjoner og er for perioden 1999 til 2019. Vi har også beregnet nye variabler under hver for der det er hensiktsmessig for bedre å kunne gjøre en god analyse, derav logaritmisk skala.

Hvorfor har vi disse med og hvordan kan de påvirke den avhengige variabelen i Solow-modellen?

Gdpcc altså BNP per innbygger er med fordi den er viktig med tanke på Solow-modellen, den gir uttrykk for produksjonen og over tid vise veksten i landet pr arbeider, og at det er et mål på velstand og ytelse. I Solow-modellen er denne den avhengige variabelen (som i modellen heter y). I datasettet har vi også med snittet for årlig vekstrate for BNP per innbygger, samt beregnet gjennomsnittlig BNP. Log BNP er logaritmen av BNP per innbygger for enklere bruk av plots videre i oppgaven.

Hci_indexen måler nivået på utdanning og ferdigheter bland befolkningen, her antar vi med Solow-modellen at hele befolkningen regnes som arbeidskraft. Ved økning i humankapital, eller kvaliteten på arbeiderne, kan dette bidra til økning i produksjon (y) og veksten på lang sikt.

Nry er naturressursene, og ved å sikre tilgangen til disse ved bærekraftig bruk, så vil denne påvirke den langsiktige veksten ved å ikke være for stor negativ.

Savings er spareraten for landet, den sier noe om hvor mye som investeres og spares og ved økning i denne, så påvirkes veksten i positiv retning.

Popgrowt er befolknings-vekstraten som sier noe om hvor mye befolkningen vokser. Denne variabelen har innvirkning på både sparing og tilbudet av arbeidskraft. Sparing fordi den kan øke eller senke sparingen ved økning eller reduksjon i befolkningsveksten, gitt at sparingen er konstant. Ved befolkningsvekst må det spares og investeres mer for å beholde steady-state, eller for å unngå reduksjon i landets vekst.

Alle disse er viktige for å gi et bilde på hvordan eller hva som påvirker veksten i et land og kunne sammenlignes med andre land basert på de samme målene. Ved å bruke Solow-modellen i en empirisk analyse ved bruk av OLS (minste-kvadrats-metode) kan finne ut om det finnes en sammenheng mellom variablene. I de neste to oppgavene skal vi se om det finnes sammenheng mellom spareraten og BNP per innbygger, befolkningsvekst og BNP per innbygger, humankapital og BNP per innbygger, forbruket av naturressurser og vekstraten i BNP per innbygger samt initialt nivå på BNP per innbygger og vekstraten i BNP per innbygger.

I tabellen under kan vi se en oversikt med spredningsmål over de variablene vi har lagt til grunn for videre analyse. Den inneholder gjennomsnittet, maksimumsverdi, minimumsverdi samt standardavviket.

	Gjennomsnitt	Standardavvik	Min verdi	Maks verdi
Gj. BNP	21533.2	22231.14	753.4	120880.6
Gj. årlig vekstrate	1.73	2.39	-10.34	7.77
BNP per innbygger	15848.6	18718.15	631.8	95684.3
Log BNP per innbygger	9	1.21	6.44	11.46
Sparerate	10.38	10.13	-19.37	45.65
Befolkningsvekstrate	1.3	1.25	-2.39	4.74
Humankapital	0.56	0.14	0.29	0.88
Humankapital kvinner	0.58	0.15	0.28	0.89
Humankapital menn	0.55	0.14	0.29	0.87
Naturressurser	4.8	7.16	0	38.6

Tabell 1: Spredningsmål

Utfordring 2.2

Empirisk analyse og test av prediksjonene fra den grunnleggende Solow-modellen.

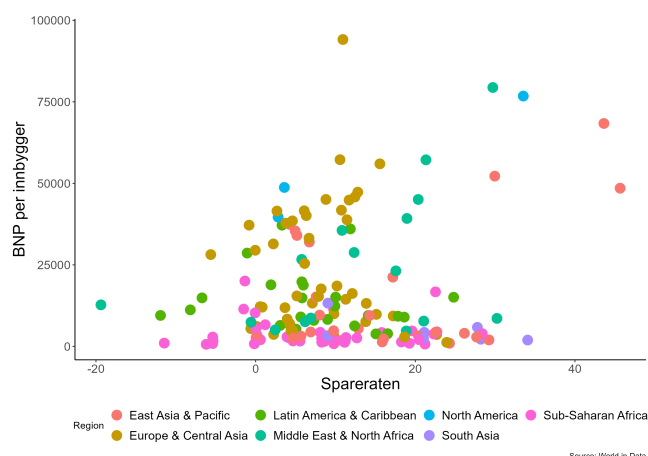
a. Lag spredningsdiagram (scatterplots) som illustrerer sammenhengen mellom spareraten med BNP per innbygger og befolkningsvekstraten med BNP per innbygger.

I følge teorien fra utfordring 1 så predikeres det i Solow-modellen at når spareraten øker så vil BNP per innbygger øke. Dette er fordi økt sparing fører til økt investering som igjen fører til økt kapital og økt produksjon.

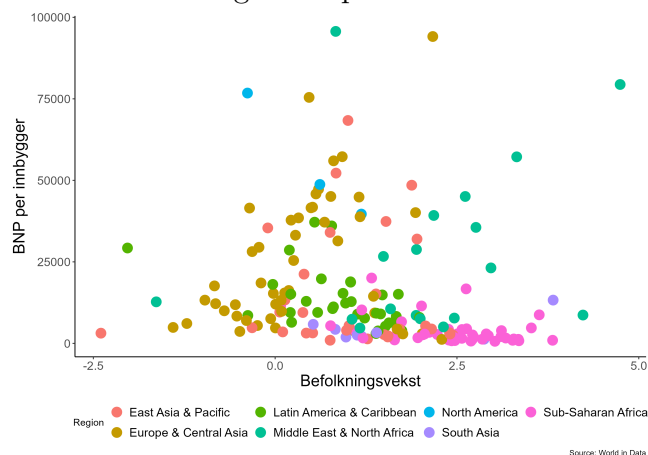
I Figur 1 ser vi at det er en svak positiv sammenheng mellom spareraten og BNP per innbygger. Dette betyr at land med høyere sparerate har høyere BNP per innbygger. Dette er i tråd med prediksjonene fra Solow-modellen, som sier at økt sparing fører til økt investering, som igjen fører til økt kapital og økt produksjon.

I Figur 2 ser vi at det er en svak negativ sammenheng mellom befolkningsvekst og BNP per innbygger. Dette betyr at land med høyere befolkningsvekst har lavere BNP per innbygger. Dette er også i tråd med prediksjonene fra Solow-modellen, som sier at økt befolkningsvekst fører til lavere BNP per innbygger, fordi det blir flere mennesker som skal dele på investeringene. Hvis befolkningen vokser raskere, så trengs det mer kapital per arbeider, som gjør at produksjon per arbeider går ned.¹

Det er svake sammenhenger vi kan observere i figurene så derfor estimerer vi en OLS (minste kvadrats-metode) som tester om spareraten og befolkningsveksten forklarer variasjon i BNP per innbygger mellom ulike land.



Figur 1: Sparerate



Figur 2: Befolkningsvekst

¹Sitat hentet fra utfordring 1 i SOK-2011 fra kandidatnummer 70, 84, 72

b. Estimer en regresjonsmodell (minste kvadrats-metode) som tester om spareraten og befolkningsveksten forklarer variasjon i BNP per innbygger mellom ulike land.

Hvorfor kjører vi en regresjonsanalyse? Regresjonsanalyse hjelper oss med å forstå hvordan ulike variabler henger sammen. Den brukes til å forutsi verdier av en avhengig variabel basert på en eller flere uavhengige variabler. Den er effektiv til å analysere store datamengder for å finne trender og mønstre som kanskje ikke er tydelige ved første øyekast.

Modellen vårs skal teste om BNP per innbygger kan forklares av spareraten og befolkningsveksten. Modellen er som følger:

$$y_{i,2015-2019} = \alpha_y + \beta_1 \cdot s_{i,2010-2015} + \beta_2 \cdot n_{i,2010-2015} + \epsilon_i$$

Hvor $y_{i,2015-2019}$ er BNP per innbygger i land i i perioden 2015-2019, $s_{i,2010-2015}$ er spareraten i land i i perioden 2010-2015, $n_{i,2010-2015}$ er befolkningsveksten i land i i perioden 2010-2015, α_y er en konstant, β_1 og β_2 er koeffisienter som forteller oss hvor mye BNP per innbygger endrer seg når spareraten og befolkningsveksten endrer seg, og ϵ_i er et feilledd.

Tabell 2: Regresjonsmodel sparerate og befolkningsvekst påvirkning på bnp

	Avhengig variabel
	BNP per innbygger
Sparerate	570.892*** (172.252)
Befolkningsvekst	-4,255.577*** (1,419.727)
Constant	21,623.470*** (2,789.207)
Observations	165
R ²	0.090
Adjusted R ²	0.079
Residual Std. Error	21,684.770 (df = 162)
F Statistic	8.037*** (df = 2; 162)
Note:	*p<0.1; **p<0.05; ***p<0.01

På neste side tolker vi resultatene fra regresjonsanalysen.

c. Tolke resultatene fra spredningsdiagrammen og regresjonsanalysen.

Constanten eller α_y i regresjonsanalysen er 21623.47. Dette betyr at hvis spareraten og befolkningsveksten er 0, så vil BNP per innbygger være 21623.47. Dette tallet gir oss en pekepinn på hvor mye BNP per innbygger ville vært hvis spareraten og befolkningsveksten var 0.

Koeffisienten til spareraten eller β_1 er positiv. Dette betyr at hvis spareraten øker så vil BNP per innbygger øke. Dette er i tråd med prediksjonene fra Solow-modellen og *** sier at resultatet er signifikant og ikke har oppstått ved tilfeldighet ($p < 0.05$).

Koeffisienten til befolkningsveksten eller β_2 er negativ. Dette betyr at hvis befolkningsveksten øker, så vil BNP per innbygger gå ned. Dette er også i tråd med prediksjonene fra Solow-modellen og *** sier at resultatet er signifikant og ikke har oppstått ved tilfeldighet ($p < 0.05$) her også, dette blir forklart bedre hvordan fungerer i [del 2.3](#).

Degrees of freedom (observations) er antall uavhengige observasjoner i datasettet minus antall parametere som estimeres i modellen. I dette tilfellet er observations 165, som betyr at det er 165 uavhengige observasjoner i datasettet, noe som da er antall land som er med i datasettet.

Residual standard error er et mål på hvor mye de avhengige variablene varierer rundt den estimerte regresjonslinjen. I dette tilfellet er residual standard error 21684.7, som betyr at de avhengige variablene varierer rundt den estimerte regresjonslinjen med 21684.7 enheter, som kan være en svakhet med modellen.

R^2 er et mål på hvor mye variasjon i den avhengige variabelen som forklares av de uavhengige variablene. I dette tilfellet er R^2 0.079, som betyr at omtrent 8% av variasjonen i BNP per innbygger kan forklares av spareraten og befolkningsveksten. Dette er et relativt høyt tall for BNP per innbygger, selv om det er mye av variansen som ikke kan forklares av modellen.

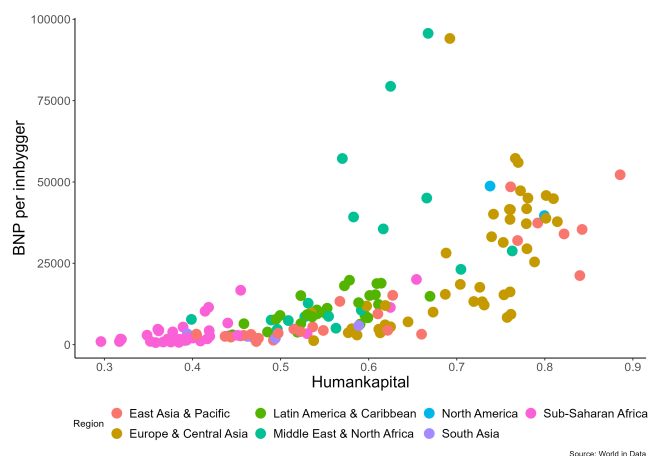
Totalt sett så er resultatene fra regresjonsanalysen i tråd med de små sammenhengene man kan se fra [Figur 1](#) og [Figur 2](#) og likt som hva Solow-modellen predikerer og viser at spareraten og befolkningsveksten kan forklare variasjon i BNP per innbygger mellom ulike land.

Utfordring 2.3

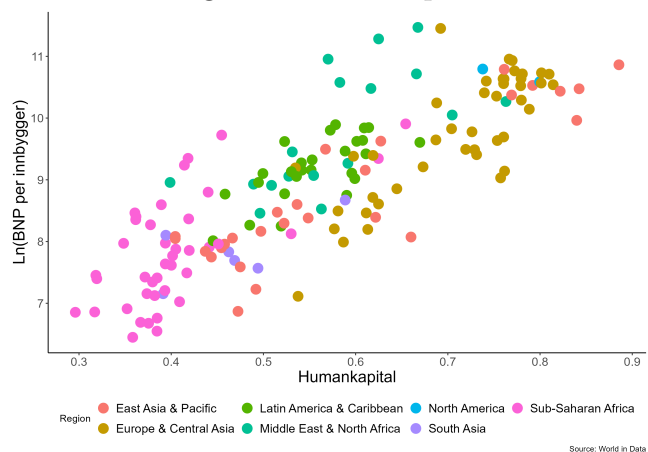
a. Spredningsdiagram som viser sammenhengen mellom Humankapital, forbruket av naturressurser og og initialt nivå på BNP per innbygger mot vekstraten i BNP per innbygger.

I motsetning til figurer i del 2.2 så kan vi i Figur 3 se at humankapital har en påvirkning på BNP per innbygger. I Europa og Sentral-Asia ser vi spesielt en vekst i BNP når humankapital øker, mens vi ser at Afrika sør for Sahara har en lavere BNP per innbygger og lavere humankapital. Dette er i tråd med Solow-modellen, som sier at humankapital er en viktig faktor for økonomisk vekst, men i motsetning til Utfordring 1, så ser det ikke ut til å være en lineær sammenheng mellom humankapital og BNP per innbygger.

Dersom vi tar humankapital i mot logaritmen til BNP per innbygger vist i Figur 4 så kan vi se et langt mye tydeligere bilde av sammenhengen mellom humankapital og BNP per innbygger. Vi ser at det er en positiv sammenheng mellom humankapital og BNP per innbygger i alle regioner i verden. Dette er i tråd med Solow-modellen. Når vi gjør lineær regresjonsanalyse så blir vi å bruke logaritmen av bnp per innbygger siden uten logaritmen så er det en sammenheng men den er ikke lineær.



Figur 3: Humankapital

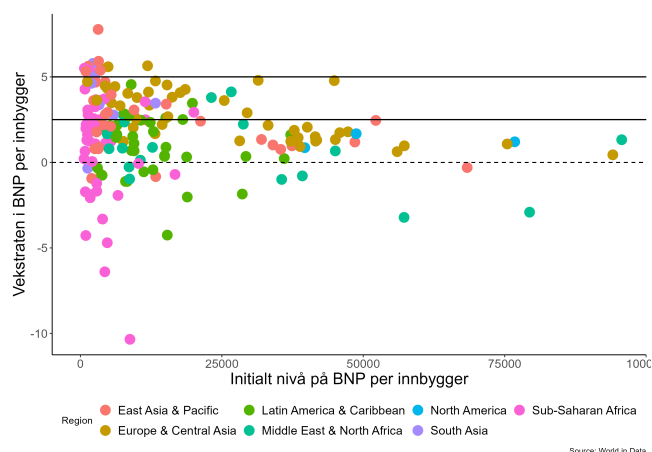


Figur 4: Humankapital logaritmisk BNP per innbygger

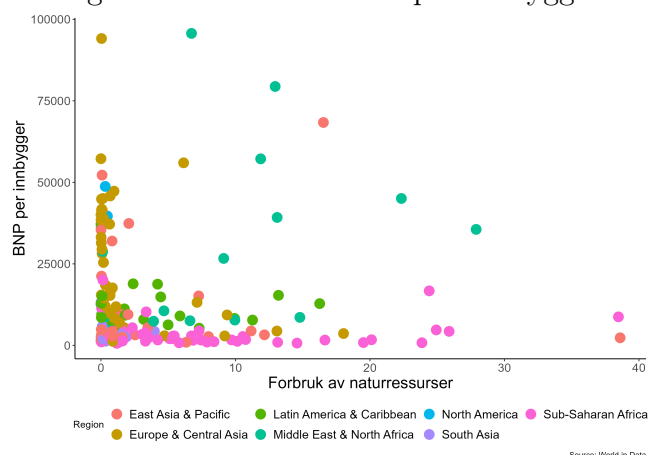
I Figur 5 ser vi effekten det initielle nivået på BNP i perioden hadde på vekstraten i BNP per innbygger. Det er tegnet inn en stripet linje på 0% vekst, med solide linjer på 2.5% og på 5%. Her ser vi at det er veldig få land som har en veldig høy BNP per innbygger som også har en høy vekstrate, men de fleste som har den høye vekstraten har et lavt initialt nivå på BNP per innbygger.

Figur 6 viser forbruk av naturressurser sin effekt på i BNP per innbygger. Her er det noe vanskeligere å se noe lineær sammenheng, men det ser ut til at land som har et høyt forbruk av naturressurser har en lavere BNP per innbygger hvor vi kan se en samling langs x-aksen. Langs y-aksen nær 0 forbruk i naturressurser ser vi en gruppering av land med høy BNP per innbygger. Denne L formen vi får på figuren kan fortelle oss at det er en negativ sammenheng mellom forbruk av naturressurser og BNP per innbygger.

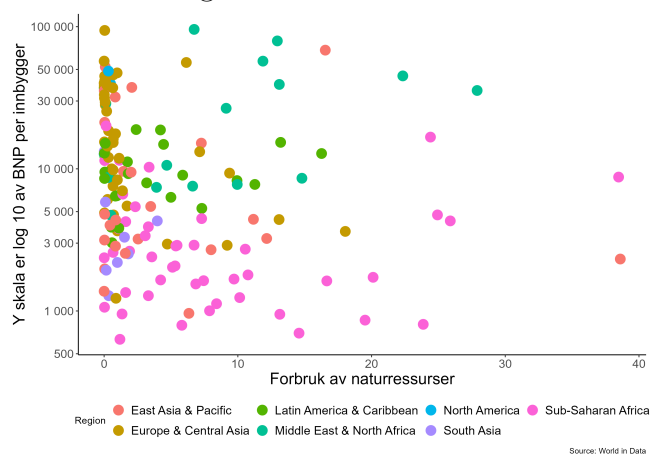
Ved å ta å skalere y-aksen med \log_{10} i Figur 7 så blir det en noe tydeligere sammenheng men **merk** at det er en annen skala på y-aksen i denne figuren. Det vi nå kan se er at de fleste landene som har under 10000 i BNP per innbygger har et høyere forbruk av naturressurser. Vi ser også at dette i stor grad er afrikanske land der det kan være andre faktorer som spiller inn som foresaker deres lavere BNP per innbygger. Men samtidig er det en samling av land med høy BNP per innbygger som har et lavt forbruk av naturressurser.



Figur 5: Vekstrate i BNP per innbygger



Figur 6: Naturressurser



Figur 7: Naturressurser med log y-akse

b. Estimer en regressionsmodell (minste kvadrats-metode) som tester om spareraten, befolkningsvekstraten, humankapital, forbruk av naturressurser og initialt nivå på BNP per innbygger, forklarer variasjon i vekstraten i BNP per innbygger i ulike land. Den modellen dere skal estimere kan bli beskrevet av ligningen under (årsaken til at vi bruker logaritmen av y_0 er fremst at BNP per innbygger er mye større enn de andre variablene. Ved å bruke den naturlige logaritmen får vi en bedre “fit” av data til modellen:

$$g_{y,i,2015-2019} = \alpha_g + \delta_1 \cdot s_{i,2010-2015} + \delta_2 \cdot n_{i,2010-2015} + \delta_3 \cdot m_{i,2015-2019} + \delta_4 \cdot u_{i,2010-2015} + \delta_5 \cdot \ln(y_0) + \epsilon_i$$

```
#Estimering av regresjonsmodell med og uten logaritme
mreg_gdp <- lm(gy ~ s + n + u + hci + ln_gdppc0, data = data)
mreg_gdp2 <- lm(gy ~ s + n + u + hci + gdppc0, data = data)
```

Tabell 3: Multipl linear regresjonsmodeller med log og ikke log avBNP per Innbygger

	Avhengig variabel	
	Vekstrate i BNP per innbygger	
	(1)	(2)
Sparerate	0.032* (0.016)	0.036** (0.018)
Befolkningsvekst	-0.628*** (0.170)	-0.524** (0.203)
Forbruk av naturressurser	-0.074*** (0.025)	-0.080*** (0.027)
Humankapital	6.901*** (2.128)	2.629 (2.169)
Initialt nivå på BNP per innbygger (log)	-1.278*** (0.224)	
Initialt nivå på BNP per innbygger		-0.00005*** (0.00001)
Constant	10.242*** (1.300)	1.813 (1.262)
Observations	139	139
R ²	0.404	0.313
Adjusted R ²	0.381	0.287
Residual Std. Error (df = 133)	1.745	1.874
F Statistic (df = 5; 133)	18.018***	12.097***

Note:

*p<0.1; **p<0.05; ***p<0.01

c. Tolke resultatene fra spredningsdiagrammene og regresjonsanalysen, og diskutere eventuelle svakheter eller begrensninger.

Stjernene i tabellene er en indikasjon på signifikansnivået, hvor * betyr at koeffesienten er signifikant på 10% signifikansnivå, ** betyr at koeffesienten er signifikant på 5% signifikansnivå, og *** betyr at koeffesienten er signifikant på under 1% signifikansnivå. Så når vi sier at noe er signifikant på et 10% signifikansnivå (vist med * og $p < 0.1$), betyr dette at det 90% sannsynlighet eller er mindre enn 10% sannsynlighet for at den observerte effekten eller en mer ekstrem effekt, kunne ha oppstått gjennom tilfeldighet alene, under forutsetning av at nullhypotesen er sann.

I begge modellene har spareraten en positiv effekt på vekstraten i BNP per innbygger. Med p-verdier indikert som * og **, er effekten statistisk signifikant, selv om styrken på signifikansen varierer mellom modellene. Dette betyr at en økning i spareraten kan forventes å øke BNP per innbygger, noe som støtter solow modellens prediksjon om spareratens betydning for økonomisk vekst.

Befolkningsveksten har en negativ effekt på BNP per innbyggers vekstrate i begge modellene, med en høy grad av statistisk signifikans (*** og **). Dette er i tråd med solow modellens prediksjoner, som antyder at høyere befolkningsvekst gjør at kapital må deles på flere personer, noe som kan redusere BNP per innbygger.

Den negative koeffisienten for forbruk av naturressurser, som er statistisk signifikant på 1% nivå i begge modeller, antyder at intensivt forbruk av naturressurser kan ha en negativ innvirkning på økonomisk vekst. I solow modellen er det en begrenset mengde naturressurser, og overforbruk kan derfor hemme bærekraftig vekst noe passer bra som en forklaring på denne sammenhengen.

Humankapital har en svært positiv effekt på BNP per innbyggers vekstrate i modell (1), med en statistisk signifikans på 1% nivå. I modell (2) er effekten fortsatt positiv, men ikke statistisk signifikant. Som vi så i [Figur 3](#) og [Figur 4](#) så var det ikke en lineær sammenheng mellom humankapital og BNP per innbygger men dersom vi tok logaritmen av BNP per innbygger så ser vi en lineær sammenheng. Dette kan være en forklaring på hvorfor humankapital ikke er signifikant i modell (2). Da vi går ut ifra modell (1) så passer dette til solow modellens prediksjon om at humankapital positivt påvirker økonomisk vekst.

Den negative koeffisienten for det initielle nivået på BNP per innbygger (både i log og ikke-log form) i hver modell indikerer en konvergensprosess, hvor land med lavere initialt BNP per innbygger vokser raskere enn de med høyere initialt BNP per innbygger. Dette er i tråd med konvergensteorien der kapital akkumuleres raskere i land med lavere BNP per innbygger, og langsommere i land med høyere BNP per innbygger.

Konstantleddet er forklart i [del 2.2](#).

Begge modellene har relativt høy Adjusted R^2 , med modell (1) som forklarer 38.1% og modell (2) som forklarer 28.7% av variansen i vekstraten i BNP per innbygger. Med så høy forklaringsgrad så kan vi anta at modellen passer godt selv om det fortsatt er en betydelig mengde varians som ikke forklares av modellene.

F-statistikken og RSE går vi ikke så mye inn på her.

Modellene antar lineære sammenhenger men som vi så i [Figur 3](#) så var det ikke en lineær sammenheng mellom humankapital og BNP per innbygger og dette er noe som også kan være et problem med de andre variablene våre.

Ekstra. Vi tester antagelsene i lineær regresjon med logaritme av intial BNP per innbygger

Kollinearitet

Vi tester kollinearitet for å se om det er korrelasjon mellom de uavhengige variablene siden dette kan kunstig øke svingingene i modellen og svekker nøyaktigheten til modellen slik at vi ikke kan stole på p verdiene.

Vi gjør nå en Variance Inflation Factor test for å få en målestokk for graden av kollinearitet.

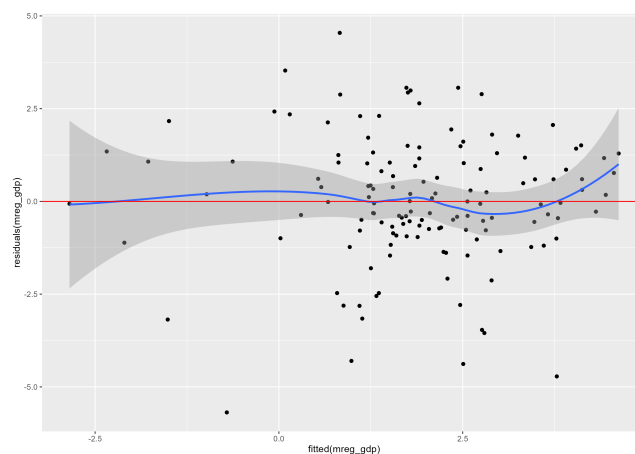
s	n	u	hci	ln_gdppc0
1.125082	1.976282	1.285129	4.483703	3.368604

Nei å nei. Moderat nesten høy kollinearitet mellom humankapital og ln_gdppc0

Linearitet

Vi setter opp en figur med residualene på yaksen og de predikerte verdiene på xaksen.

Holder seg noe bra men avviker på slutt.



Figur 8: Linearitetstest

Normalitet

Vi sjekker normaliteten i et kvantil-kvantil plot.

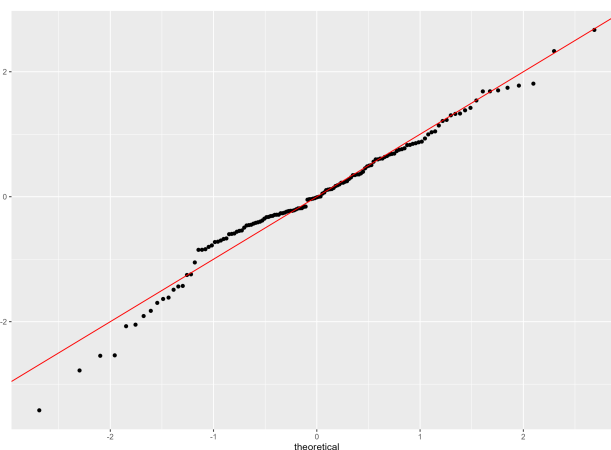
Huff

Homoskedasitet

Vi lager en figur der jeg tar kvadratroten av absoluttverdiene til de standardiserte residualene på yaksen og de predikerte verdiene på xaksen.

Kan ikke brukes som linjal.

Gjør nå en Non Constant Variance test



Figur 9: Normalitetstest

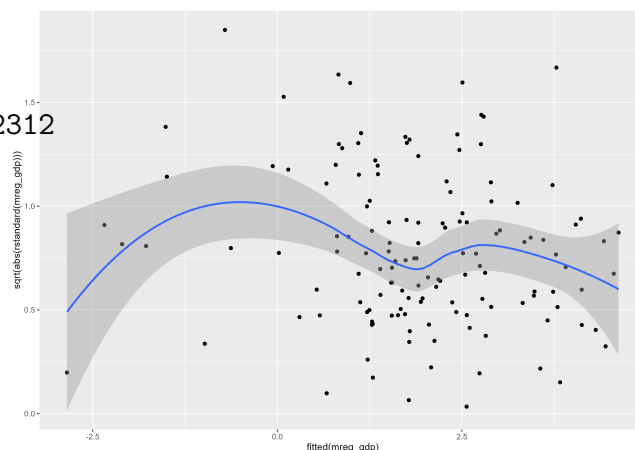
Non-constant Variance Score Test

Variance formula: $\sim \text{fitted.values}$

Chisquare = 5.221321, Df = 1, p = 0.022312

Heteroskedastisk?

Modellen har brudd, så dessverre kan vi ikke si at det vi har sett på har noen statistisk sammenheng.



Figur 10: Homoskedasitetstest

Referanser

- DataBank, the world bank.* (2024). <https://databank.worldbank.org/home.aspx>
- Hess, P. N. (2016). *Economic growth and sustainable development*. Routledge.
- Hlavac, M. (2022). *Stargazer: Well-formatted regression and summary statistics tables*. Social Policy Institute. <https://CRAN.R-project.org/package=stargazer>
- Mannberg, A. (2023). *SOK-2011 v2024: Utfodring 2*. https://uit-sok-2011-v2024.github.io/assets/sok2011_utf2_2024.html

Appendix Generell KI bruk

I løpet av koden så kan det ses mange # kommentarer der det er skrevet for eks “#fillbetween q1 and q2”. Når vi skriver kode i Visual Studio Code så finnes det en plugin som heter Github Copilot. Når vi skriver slike kommentarer så kan den foresøke å fullføre kodelinjene mens vi skriver de. Noen ganger klarer den det, men andre ikke. Det er vanskelig å dokumentere hvert bruk der den er brukt siden det “går veldig fort” men siden det ikke er fått på plass en slik dokumentasjon så kan all python kode der det er brukt kommentarer antas som at det er brukt Github Copilot. Nærmere info om dette KI verktøyet kan ses på <https://github.com/features/copilot>

Vi har brukt KI til å skrive sammendrag og kvalitetssikret at det som er skrevet er korrekt.

Appendix 1.1 KI bruk

Kandnummer 84 har brukt KI til å få en oversikt på hvordan legge opp skrivingen og noe struktur

Analyse av økonomisk vekst

April 4, 2024



Anonymous

Har fått denne oppgaven i samfunnsøkonomi med økonomisk vekst og bærekraft, hvordan bør jeg sette oppgaven opp og hva bør jeg har med? Beskriv data-materialet som du vil bruke i analysen. Beskrivelsen skal inneholde en beskrivelse av datakilden, og variablene som blir brukt. Forklar hvorfor du har med disse variablene. Lag en tabell som viser gjennomsnittsverdier, standard-avvik, min- og maks-verdi for alle variabler.



ChatGPT

Når du skal skrive oppgaven din om samfunnsøkonomi med fokus på økonomisk vekst og bærekraft, er det viktig å følge en struktur som gjør det mulig for leseren å forstå din tilnærming og analyse. Her er en anbefalt struktur for oppgaven din:

- Introduksjon:
 - Gi en kort oversikt over temaet og hvorfor det er viktig å analysere økonomisk vekst med tanke på bærekraft.
 - Presenter problemstillingen eller forskningsspørsmålet ditt.
- Teoretisk rammeverk:
 - Gi en kort gjennomgang av teoriene og konseptene knyttet til økonomisk vekst og bærekraft.
 - Forklar hvordan disse to konseptene kan være i konflikt eller samarbeide.
- Metode:
 - Beskriv hvilke data du vil bruke og hvorfor du har valgt disse dataene.
 - Forklar hvilke variabler du vil inkludere i analysen, og hvorfor disse variablene er relevante.
 - Gi en kort beskrivelse av metoden du vil bruke for å analysere dataene, for eksempel regresjonsanalyse eller deskriptiv statistikk.
- Data:
 - Beskriv datakilden din, for eksempel en datasett fra en offentlig databank, en spørreundersøkelse eller tidsseriedata.
 - List opp de ulike variablene i datasettet og gi en kort forklaring på hver variabel.
 - Forklar hvorfor hver variabel er relevant for din analyse av økonomisk vekst og bærekraft.
- Analysen:
 - Utfør analysen av dataene ved hjelp av den valgte metoden.
 - Presenter resultatene dine på en klar og sammenhengende måte.
 - Diskuter funnene dine i forhold til problemstillingen eller forskningsspørsmålet ditt.
- Diskusjon og konklusjon:
 - Diskuter implikasjonene av dine funn for økonomisk politikk og praksis.
 - Gi en oppsummering av hovedfunnene dine og eventuelle begrensninger ved studien.
 - Gi forslag til videre forskning på området.

Når det gjelder data-materialet ditt, kan du for eksempel bruke økonomiske indikatorer som BNP-vekst, arbeidsledighet, inflasjon og miljøindikatorer som karbonutslipp, energiforbruk og bruk av fornybare ressurser. Du kan også vurdere å inkludere variabler som investeringer i forskning og utvikling, utdanningsnivå, og politiske tiltak rettet mot bærekraftig utvikling.

Her er en tabell som viser gjennomsnittsverdier, standardavvik, min- og maksverdier for noen hypotetiske variabler:

Variabel	Gjennomsnitt	Standardavvik	Min. verdi	Maks. verdi
BNP-vekst (%)	2.5	1.2	0.8	4.6
Arbeidsledighet (%)	5.0	1.8	2.2	8.7
Karbonutslipp (tonn)	5000	1200	3200	7800
Andel fornybar energi (%)	30	10	20	45

Denne tabellen gir leseren en rask oversikt over variasjonen og gjennomsnittet av de ulike variablene som er brukt i analysen din.



Anonymous

innenfor data delen, hvordan bør jeg legge denne opp?



ChatGPT

Når du beskriver dataene du skal bruke i analysen din, er det viktig å være detaljert og presis. Her er en strukturert tilnærming du kan følge for å legge opp data-delen av oppgaven din:

- Beskrivelse av datakilden:
 - Start med å identifisere og beskrive kilden til dataene dine. Dette kan være en offentlig databank, et datasett fra en forskningsinstitusjon, eller andre kilder.
 - Gi en kort historisk kontekst for datakilden, inkludert hvordan dataene ble samlet inn, hyppigheten av oppdateringer og eventuelle metodologiske begrensninger.
- Variabler:
 - List opp de ulike variablene som er inkludert i datasettet ditt.
 - Gi en kort forklaring på hver variabel, inkludert enheter og måleenheter.
 - Forklar hvorfor hver variabel er relevant for analysen din av økonomisk vekst og bærekraft.
- Datainnsamling og behandling:
 - Beskriv hvordan dataene ble samlet inn, og eventuelle metoder som ble brukt for å rense og behandle dataene.
 - Diskuter eventuelle utfordringer eller begrensninger knyttet til datainnsamling og behandling.
- Statistisk analyse:
 - Gi en oversikt over hvilke statistiske metoder du planlegger å bruke for å analysere dataene dine, for eksempel deskriptiv statistikk, regresjonsanalyse eller korrelasjonsanalyse.
 - Forklar hvorfor hver metode er egnet for analysen din og hvordan den vil bidra til å besvare problemstillingen din.
- Kvalitetssikring:
 - Diskuter tiltak du har tatt for å sikre kvaliteten på dataene dine, for eksempel validering av datasettet, identifisering og håndtering av outliers, og sensitivity testing.
- Beskrivelse av datatabell:
 - Inkluder en tabell som viser gjennomsnittsverdier, standardavvik, min- og maksverdier for alle variabler, som jeg viste i forrige svar.
 - Dette gir leseren en rask oversikt over variasjonen i datasettet ditt.

Ved å følge denne strukturen vil du kunne gi en grundig og sammenhengende beskrivelse av dataene dine og hvordan de vil bli brukt i analysen din av økonomisk vekst og bærekraft.

Is this conversation helpful so far?

Get started with ChatGPT

[Report content](#) | [Terms of use](#) | [Privacy policy](#)

Appendix Kapittel 1.2 KI bruk

Ingenting

Appendix Kapittel 1.3 KI bruk

Ingenting