



Den lange veien fra samvariasjon til årsakssammenheng

Separating causation from correlation

Ola Lotherington Vestad
Forsker, Statistisk sentralbyrå
Ola.Vestad@ssb.no

Hovedformålet med artikkelen «Immigration and Social Mobility» er å forsøke å besvare følgende spørsmål: Kan (veksten i) innvandring til Norge de siste tiårene være forklaringen på hvorfor personer med bakgrunn i lavere sosiale klasser har sakkert akterut på arbeidsmarkedet, relativt til personer med bakgrunn i høyere sosiale klasser? Forfatterne studerer hvordan forskjeller mellom personer med høy og lav sosioøkonomisk klassebakgrunn har endret seg over tid, og hvordan endringer i disse forskjellene varierer mellom regioner med høy og lav innvandrerandel. Resultatene tyder på at personer fra lavere sosiale klasser gjør det dårligere på arbeidsmarkedet, relativt til personer fra høyere sosiale klasser, i regioner med høy innvandring fra lavinntektsland enn i regioner med lav innvandring fra lavinntektsland. Samtidig ser det ut til at forskjellene mellom høy og lav sosioøkonomisk klassebakgrunn har blitt mindre i regioner med høy innvandring fra høyinntektsland, sammenliknet med regioner med lav innvandring fra høyinntektsland.

Forfatterne studerer sammenhengen mellom to interessante trekk ved den norske samfunnsutviklingen i nyere tid. Vi har sett en sterk vekst i innvandring, særlig fra lavinntektsland, som har vært ulikt fordelt mellom ulike geografiske regioner. Samtidig ser vi også noen tegn til redusert økonomisk og sosial mobilitet. Solid empirisk forskning som kan bidra til økt forståelse av hvordan innvandring påvirker samfunnsutviklingen generelt, og økonomisk mobilitet spesielt, vil utvilsomt være til stor nytte for den offentlige debatten og for politikktutforming i Norge. Samtidig vil erfaringene fra Norge kunne være relevante for andre land hvor utviklingen i innvandring og økonomisk mobilitet har gått i samme retning som i Norge.

Artikkelen inneholder mye interessant empiri om økonomisk mobilitet og innvandring til Norge i perioden fra 1992 til 2016. Både veksten i innvandring i perioden og den store spredningen i innvandringsandeler på tvers av regioner er godt dokumentert, og dermed er det lett å gå med på at sammenlikninger mellom regioner med høy og lav vekst i innvandring kan være både nyttig og relevant. Forfatterne gjør også en god jobb med å begrunne skillet mel-

lom innvandring fra lav- og høyinntektsland. For det første viser de at veksten i innvandring nesten utelukkende er drevet av en vekst i innvandring fra lavinntektsland. For det andre viser de at det er store forskjeller mellom innvandrere fra lav- og høyinntektsland i form av hvilke jobber de får: innvandrere fra lavinntektsland ser ut til å konkurrere om de samme jobbene som norskfødte fra lavere sosiale klasser, mens innvandrere fra høyinntektsland er overrepresentert i yrker som utføres av norskfødte fra høyere sosiale klasser.

Men samtidig som den empiriske spesifikasjonen er godt begrunnet og forarbeidet er godt utført, har artikkelen vesentlige mangler når det kommer til å skille mellom samvariasjon og årsakssammenhenger. Mitt viktigste ankepunkt mot en kausal tolkning av resultatene er ganske enkelt at forfatterne ikke gjør en god nok jobb med å sannsynliggjøre at den identifiserende antakelsen holder. Nedenfor har jeg gjort et forsøk på å utdype hvorfor jeg mener at den identifiserende antakelsen ikke er tilstrekkelig godt begrunnet, etterfulgt av et forslag til hvordan problemet kan løses.

DEN IDENTIFISERENDE ANTAKELSEN MÅ SANNSYNLIGGJØRES

Forfatterne tar utgangspunkt i en variant av en «triple-differences»-spesifikasjon hvor effekter av innvandring på ulike arbeidsmarkedsutfall for norskfødte identifiseres av tre forskjeller: forskjeller mellom regioner med høy og lav innvandring, forskjeller over tid (før og etter veksten i innvandring) og forskjeller mellom personer med høy og lav sosioøkonomisk klassebakgrunn. For å identifisere effekter av innvandring i dette rammeverket, må en legge til grunn en antakelse om at regioner med høy og lav vekst i innvandring hadde hatt lik utvikling i forskjeller mellom personer med høy og lav klassebakgrunn hvis det ikke hadde vært for forskjeller i veksten i innvandring. Hvis denne antakelsen holder kan vi bruke den faktiske utviklingen i forskjeller mellom personer med høy og lav klassebakgrunn i regioner med lav vekst i innvandring som en approksimasjon til den kontrafaktiske utviklingen i forskjeller mellom personer med høy og lav innvandring i regioner med høy vekst i innvandring. Vi antar altså at klasseforskjellene i regioner med høy vekst i innvandring hadde utviklet seg på samme måte som klasseforskjellene i regioner med lav vekst i innvandring hvis det ikke hadde vært for at veksten i innvandring var ulikt fordelt mellom regioner.

Denne identifiserende antakelsen er ikke bare vanskelig å formulere, det er også vanskelig å vurdere om den holder eller ikke. Er det for eksempel rimelig å legge til grunn at den teknologiske utviklingen (eller andre relevante endringer i perioden som studeres) har hatt lik effekt på forskjeller mellom høy og lav klasse i regioner med høy vekst i innvandring som i regioner med lav vekst i innvandring? Ingen kan vite helt sikkert om en identifiserende antakelse av denne typen holder eller ikke. Derfor må vi som forskere argumentere for og imot, og vi må gjøre det vi kan for å fremskaffe empiri som gjør det mulig å vurdere gyldigheten av den identifiserende antakelsen, selv om den ikke lar seg teste direkte. Det er særlig på dette området jeg mener at Hoen et al. kunne og burde ha gjort en grundigere jobb.

Den mest sentrale indirekte testen for identifikasjonsstrategier basert på forskjeller-i-forskjeller («difference-in-differences»), er relatert til endringer i utfall før «treatment» inntreffer. Og i dette konkrete tilfellet handler det om å undersøke om forskjellene mellom høy og lav klasse fulgte den samme utviklingen i regioner med høy vekst i innvandring som i regioner med lav vekst i innvandring, før veksten i innvandring startet (eller tiltok). Var

det noen tegn til økende sosiale forskjeller før veksten i innvandring tiltok? Jo mer ulik utvikling i forskjeller mellom høy og lav klasse for regioner med høy og lav vekst i innvandring før veksten i innvandring, desto mindre sannsynlig er det at forskjellene mellom høy og lav klasse i regioner med lav vekst i innvandring er en god tilnærming til de kontrafaktiske forskjellene mellom høy og lav klasse i regioner med høy vekst i innvandring.

Det kan være flere grunner til at forfatterne ikke rapporterer noen tester av typen beskrevet ovenfor. Én mulig grunn er at forfatterne dessverre ikke har ubegrenset tilgang til data bakover i tid. Mye av datagrunnlaget strekker seg bare tilbake til 1992, og dermed blir det ikke helt enkelt å undersøke utviklingen i utfall før veksten i innvandring startet. En annen mulig grunn er at modellen som estimeres faktisk er mer generell og komplisert enn den forenklete versjonen jeg har basert meg på i avsnittene over. Modellen som estimeres skiller mellom fem ulike sosiale klasser, den skiller mellom innvandring fra lav- og høyinntektsland, og den skiller ikke bare mellom høy og lav innvandring – den bruker innvandrersandeler for hver region i hvert år. Men dette betyr ikke at det ikke er mulig å fremskaffe empiri som understøtter den identifiserende antakelsen, bare at en må ty til noen forenklinger og legge ned litt ekstra innsats for å få det til.

Én mulig fremgangsmåte er å basere seg på en forenklet versjon av modellen, i tråd med beskrivelsen over. En kunne starte med å definere to grupper av regioner: den ene gruppen hadde høy vekst i innvandring, den andre hadde lav vekst i innvandring. En kunne deretter dele tidsperioden i to: før og etter tiltakende vekst i innvandring, dvs. før og etter «treatment», og bruke dette som utgangspunkt for relevant deskriptiv statistikk. Hva kjennetegner regioner med henholdsvis høy og lav vekst i innvandring, og på hvilke områder var de to gruppene forskjellige før veksten i innvandring for alvor gjorde seg gjeldende?

Deretter ville det være nyttig å studere utviklingen i ulike utfall over tid for de to gruppene av regioner. Spesielt interessant ville det være å undersøke om regioner med høy vekst i innvandring utviklet seg noenlunde likt som regioner med lav vekst i innvandring, før veksten i innvandring tiltok. Systematiske forskjeller som avdekkes med en slik deskriptiv analyse er ikke bare interessante i seg selv, de kan også bidra med støtte til identifikasjonsstrategien. Hvis utviklingen i regioner med høy vekst i innvandring var i ferd med å avvike fra utviklingen i regioner med lav vekst i innvandring allerede før veksten i innvandring tiltok, er det lett å forstå at forfatterne ikke er komfortable med å basere seg på en standard «difference-in-differences»-identifikasjonsstrategi: da virker det ikke særlig sannsynlig at forskjellig vekst i innvandring er eneste årsak til at andre forskjeller øker mellom de to gruppene av regioner.

I så fall kan det tenkes at vi kan komme et stykke videre ved å innføre en tredje differanse: forskjeller mellom personer med ulik klassebakgrunn, eller «høy versus lav klasse». For at denne tredje differansen skal kunne løse identifikasjonsproblemet, må det altså være sånn at hva enn det er som bidro til forskjeller mellom regioner med høy og lav vekst i innvandring, før den sterke veksten i innvandring, har den samme innvirkningen på forskjeller mellom høy og lav klasse i regioner med høy vekst i innvandring som i regioner med lav vekst i innvandring. Som fagfelle er jeg ikke villig til å legge denne antakelsen til grunn før jeg har sett empiri som indikerer at den holder. Og hvis empirien tyder på at antakelsen ikke holder, må tolkningen av resultatene justeres. Da kan det ikke nødvendigvis fastslås at analysene har avdekket en kausal effekt av innvandring på sosial mobilitet, men vi kan likevel ha lært mye nyttig og nytt om både innvandring og sosial mobilitet.

Hvis empiri basert på den forenklete versjonen av modellen ikke gir grunn til å mistenke at den identifiserende antakelsen ikke holder, så er det veldig gode nyheter. Og da kan det være noe å hente på å bevege seg i retning av en mer generell modell som benytter mer av variasjonen i data, f.eks. ved å gå fra to til fem sosiale klasser og ved å gå fra ett binært til to kontinuerlige mål på innvandring.

ANDRE KOMMENTARER

Kontrollere for ulik betydning av klassebakgrunn i ulike regioner

Den empiriske spesifikasjonen kontrollerer for forskjeller mellom sosiale klasser over tid og for forskjeller mellom regioner over tid (og begge deler gjøres separat for kvinner og menn), men ikke for at forskjeller mellom klasser kan variere på tvers av regioner. Utelatelsen av interaksjonen mellom sosial klasse og region innebærer en antakelse om at en gitt sosial klassebakgrunn har samme betydning for arbeidsmarkedsutfall i alle regioner. Dette fremstår som en litt streng antakelse, særlig når to av utfallene er definert med utgangspunkt i regionale inntektsfordelinger (henholdsvis rangering i den regionale inntektsfordelingen og inntekt relativt til det regionale gjennomsnittet). Når de regionale inntektsfordelingene er forskjellige, vil det også være forskjellig hvor mye som kreves for å klatre oppover i fordelingen. Og ulik betydning av klassebakgrunn i ulike regioner vil fanges opp i innvandrerandel-koeffisientene, hvis betydningen av klassebakgrunn av en eller annen grunn er korrelert med innvandrerandel. Det ville derfor være nyttig å se om de estimerte innvandrerandel-koeffisientene endres når modellen utvides til også å inkludere interaksjonen mellom sosial klasse og region.

Regionale versus nasjonale inntektsfordelinger

Jeg ville også gjerne ha sett en mer inngående diskusjon rundt valget om å rangere foreldre og barn etter regionale inntektsfordelinger heller enn etter den nasjonale inntektsfordelingen. Med rangering etter den nasjonale inntektsfordelingen vil forskjellen mellom to desiler bety det samme for alle regioner, mens regionale rangeringer innebærer at forskjellen mellom to desiler kan bety noe helt annet i én region enn i en annen. Men samtidig som rangering nasjonalt kan være lettere å tolke enn rangering regionalt, kan f.eks. regionale forskjeller i levekostnader være et argument for å rangere regionalt. Og når det ikke er helt opplagt om man bør rangere regionalt eller nasjonalt, ville det også være nyttig å vite om resultatene endres når man går fra den ene til den andre måten å rangere på.

Standardfeil

De estimerte standardfeilene er robuste for korrelasjon mellom observasjoner innenfor samme region og år. For meg virker det nødvendig å gå ett nivå høyere for å ta høyde for korrelasjon over tid mellom observasjoner i samme region, altså cluster på region heller enn cluster på region-ganger-år.