Leif Helland

Årsak og tilfeldighet

Modeller, eksperimenter og adferd i samfunnsvitenskapene

CAPPELEN DAMM
AKADEMISK

© CAPPELEN DAMM AS, Oslo, 2021

ISBN 978-82-02-72636-2

1. utgave, 1. versjon 2021

E-utgaven baserer seg på 1. trykte utgave, 978-82-02-37930-8,
Cappelen Damm Akademisk 2016

Materialet i denne publikasjonen er omfattet av åndsverklovens bestemmelser. Uten særskilt avtale med forlaget er enhver eksemplarfremstilling og tilgjengeliggjøring bare tillatt i den utstrekning det er hjemlet i lov eller tillatt gjennom avtale med Kopinor, interesseorgan for rettighetshavere til åndsverk.

Omslagsdesign: Anders Svendsen
Sats: HAVE A BOOK
Elektronisk tilrettelegging: HAVE A BOOK

Forfatteren har mottatt støtte fra Det faglitterære fond.

www.cda.no
akademisk@cappelendamm.no

Innhold

Forord 7

KAPITTEL 1

Innledning 9
- Randomisering 11

Kontroll 16
- Validitet 30
- Oppsummering 35

Litteratur 36

KAPITTEL 2

På egen hånd 39

Individuelle beslutninger – nytteforventningsteorien 39
- Prospektteori 43
- En monetarisert apeøkonomi 49
- Tilfeldigheter og vane 54
- Gyldighetsområde og empirisk innhold 56
- Oppsummering 63

Litteratur 64

KAPITTEL 3

Forhandling 67
- Standard teori 67
- Verdifordeling i laboratoriet 82
- Tillit og tillitsbrudd 112
- Oppsummering 115

Litteratur 116

KAPITTEL 4

Dugnad 119
- Fellesgoder 119
- Teori 119
- Fellesgodeeksperimenter uten straffemuligheter 126
- Fellesgodeeksperimenter med straffemuligheter 148

Straffekulturer ... 158
Ledelse i fellesgodesituasjoner ... 162
Oppsummering ... 167
Litteratur ... 167

KAPITTEL 5
Marked ... 171
Frikonkurransemodellen ... 171
Eksperimentelle markeder: hyper-effektivitet ... 176
Bobler i et eksperimentelt finansmarked ... 193
Et marked med prisannonsering ... 198
Å kjøpe seg en amerikansk president ... 208
Oppsummering ... 212
Litteratur ... 213

KAPITTEL 6
Demokrati ... 215
Arbeidshesten ... 215
Arbeidshesten i laboratoriet ... 221
Velgerkoordinering ... 232
Velgerkoordinering i laboratoriet ... 234
Kontroll og seleksjon ... 236
Kontroll og seleksjon i laboratoriet ... 240
Lobbying ... 244
Lobbying i laboratoriet ... 247
Oppsummering ... 254
Litteratur ... 257

KAPITTEL 7
Sluttord ... 261
Metoden ... 261
Et grunnlagsspørsmål ... 261
Snevre egeninteresser ... 262
Rasjonalitet ... 264
Institusjonene ... 265
Feltet ... 266
Litteratur ... 267

Forord

Denne boken henvender seg primært til studenter i statsvitenskap og økonomi, men burde også ha interesse for studenter fra andre fagområder. Tanken har vært å gi en forholdsvis bred innføring i eksperimentell metode, og å illustrere denne ved å gå dypere inn i noen utvalgte samhandlingssituasjoner: forhandlinger, markeder, produksjon av fellesgoder, og demokratisk beslutningsfatting.

Jeg har hatt stor glede av å diskutere en rekke av problemstillingene som behandles i boken med mine gode kolleger ved Institutt for samfunnsøkonomi, Handelshøyskolen BI. En særlig takk til Erik Grønn, som tok seg bryet med å lese og kommentere tidlige utkast.

Takk til Norsk Faglitterær Forfatter – og Oversetterforening, som finansierte bokprosjektet. Takk til min forlagsredaktør Tor Paulson, for å ha øvet et mildt press nå og da for å få boken ferdigstilt. Takk til forlagets konsulent, som bidro til å forbedre fremstillingen vesentlig.

Nydalen, August 2016
Leif Helland

KAPITTEL 1

Innledning

1

Sosial samhandling står i sentrum av samfunnsvitenskapene. Ofte er slik samhandling strategisk motivert. Kandidater i valg fremmer valgløfter som de tror vil kapre velgere fra andre kandidater; bedrifter i et marked priser med sikte på å ta markedsandeler fra andre bedrifter; gruppemedlemmer bidrar til fellesprosjekter betinget av at også andre gruppemedlemmer bidrar; forhandlere betinger sitt forslag av hva de tror motparten vil kunne akseptere, og så videre.

Strategisk samhandling forstås best med spillmodeller. Slike modeller er partielle og gjennomanalyserer én enkelt mekanisme av interesse. For at vi skal kunne trekke skarpe adferdsimplikasjoner, identifiseres et velavgrenset og presist sett av forutsetninger som modellen bygger på. Dette har flere fortrinn. Først og fremst tillater det oss å sjekke om modellens forutsetninger er konsistente; i betydningen av at de kan innfries samtidig.

Fra en modell som hviler på et sett av inkonsistente (selvimotsigende) forutsetninger kan en – logisk sett – trekke den implikasjon en måtte ønske. Inkonsistente sett av forutsetninger er derfor forenlige med *alle* tenkelige observasjoner, og er uten empirisk innhold. Med et sett av konsistente forutsetninger åpner vi for at bestemte observasjoner er uforenlige med modellen: de kan altså ikke utledes fra settet av forutsetninger som modellen bygger på. Modellen har da et empirisk innhold – den er falsifiserbar.

Men hvor godt treffer modellene våre i data? I denne boken argumenteres det for at laboratorie eksperimenter er særlig velegnet til å kaste lys over nettopp dette spørsmålet.

Vi skal fokusere på eksperimenter som benytter standarder og normer utviklet innenfor det som kalles «eksperimentell økonomi». Eksperimentell økonomi er en bestemt fremgangsmåte for innsamling av data i kontrollerte omgivelser, altså en metode. Bruken av metoden er ikke begrenset til analyse av økonomiske problemstillinger, og økonomer har ikke – og ønsker heller ikke – monopol på metoden. Tvert i mot har den vist seg velegnet til å studere sentrale problemstillinger i alle deler av samfunnsvitenskapene. I denne boken ser vi på problemstillinger fra blant annet økonomi, statsvitenskap, sosiologi og adferdspsykologi.

Det er en vekselvirkning mellom eksperimentell testing av eksisterende modeller og modellutvikling. Spillmodeller bygger ofte på egeninteresse

og åpent kjent rasjonalitet. Vi skal kalle slike modeller for «standardmodeller». Hverken åpent kjent rasjonalitet eller ren egeninteresse er imidlertid nødvendige bestanddeler i en spillmodell.

Når replikerbare eksperimenter produserer adferdsmønstre som avviker fra standardmodellen, leder dette ofte til at standardmodellen justeres for å kunne forklare avvikene. Den justerte modellen brukes så til å utlede nye implikasjoner, som i sin tur testes i nye eksperimenter. På denne måten kan en håpe på progresjon i forskningen.

I de tilfeller der replikerbare eksperimenter produserer adferdsmønstre som samsvarer med standardmodellen, så er likevel ikke dette ensbetydende med at forskerne slår seg til ro med at standardmodellen forklarer adferden. Man kan gå dypere og spørre om adferden produseres av de mekanismer som standardmodellen skisserer, eller om adferden primært har andre drivkrefter. Innsikt i dette representerer også progresjon. I boken skal vi se eksempler på begge former for progresjon.

Boken har fem tematisk inndelte kapitler. I hvert av disse spør vi hvordan laboratorieeksperimenter kan kaste lys over forklaringskraften i modellene vi bruker for å forstå samfunnet rundt oss. Tematikkene som er valgt, er sentrale i samfunnsvitenskapene.

Først ser vi på grunnlaget for analysen (kapittel 2). Spørsmålet vi stiller er om individer i det hele tatt er i stand til å opptre rasjonelt.

Dernest ser vi på strategisk samhandling i situasjoner der to eller flere parter skal fremforhandle en avtale (kapittel 3). Spørsmålet er om forhandlingsadferden vi observerer i laboratoriet avspeiler de prinsipper som rasjonelle og egeninteresserte forhandlere ville fulgt for å nå frem til en avtale.

Produksjon av fellesgoder står i fokus i kapittel 4. Spørsmålet er om subjekter i laboratoriet er i stand til å overskride de sosiale dilemmaene som fellesgodeproduksjon reiser i fravær av sentralisert tvangsmakt, og hvordan de i så fall klarer det.

Prisdannelse i markeder er tema for kapittel 5. Vi spør hvor gode individer egentlig er til å oppdage likevektspriser i desentraliserte markeder, og hvilke markedsinstitusjoner som best bidrar til effektivitet.

Det siste temaet er demokrati (kapittel 6). Her er spørsmålet hvor godt velgere er i stand til å produsere representative vedtak i demokratiske institusjoner.

Boken avsluttes med et kort oppsummeringskapittel (kapittel 7).

Før vi gyver løs på bokens ulike tematikker, må vi diskutere laboratorieeksperimenter som metode. Dette er tema for resten av kapitlet. Vi tar utgangspunkt i fire stikkord:

- Modell
- Randomisering

- Kontroll
- Validitet (generaliserbarhet)

Vi argumenterer for at laboratorieeksperimenter er særlig velegnet til å avdekke generelle prinsipper, slik disse kommer til uttrykk i presise – og falsifiserbare – *modeller* av menneskelig samhandling. Det er to grunner til dette.

For det første gir laboratorieeksperimetner – i likhet med felteksperimenter – grunnlag for å identifisere gjennomsnittlige behandlingseffekter, ved at subjekter *randomiseres* inn i kontroll- og behandlingsgruppe(r). Vi kan da si noe om årsaksforhold, og er ikke henvist til å nøye oss med å påvise korrelasjon mellom variable.

For det andre gir laboratorieeksperimenter sterk grad av *kontroll*. Med kontroll mener vi evnen til å mimikere en modell i laboratoriet, slik at modellens incentivstruktur gjenskapes i en virkelig beslutningssituasjon.

Men hva kan laboratorieeksperimenter lære oss om «verden der ute»? Dette er et spørsmål om ekstern *validitet*. I boken argumenteres det for at omgivelsene som skapes i et laboratorieeksperiment er en del av en virkelig verden: hvor virkelige mennesker stilles overfor virkelige valg, med virkelige konsekvenser. Hvis eksperimentet er godt designet, kan det si noe om menneskers evne, og vilje, til å respondere på incentivstrukturen i modellene våre. Innsikt i dette setter oss bedre i stand til å velge modeller som vi har tiltro til når vi skal analysere ikke-eksperimentelle data fra naturlig forekommende beslutningssituasjoner.

Randomisering

La oss ta utgangspunkt i et enkelt eksperiment for å vise hvordan *tilfeldig* tildeling av subjekter til ulike grupper (randomisering) setter oss i stand til å trekke slutninger om kausalitet. Dette enkle knepet gir eksperimentell metode et betydelig fortrinn i forhold til andre empiriske metoder.

Ultimatumspill

Eksperimentet tar utgangspunkt i en enkel modell, ultimatumspillet. Dette spillet har to spillere, F (forslagsstiller) og M (mottaker). Før eksperimentet starter gir administrator et pengebeløp x til F. Pengebeløpet skal fordeles mellom F og M etter følgende regler: I første steg tilbyr F $s \in [0,1]$ til M, hvor s er en andel av x. I andre steg aksepterer eller forkaster M tilbudet. Hvis M aksepterer, får F $(1-s)$ og M får s. Hvis M forkaster, får de begge 0.

Det er gjennomført talløse ultimatumeksperimenter (se kapittel 3). I slike eksperimenter observeres ofte tilbud mellom 30 og 40 prosent av ressurstilgangen. Det observeres også ofte at tilbud under $20-25$ prosent av ressurstilgangen forkastes. Hvis F og M utelukkende er motivert av penger, og dessuten er rasjonelle, skulle vi ikke observere dette. I andre steg vil

en egeninteressert M akseptere alle tilbud s > 0. En egeninteressert F vil forvente dette, og derfor foreslå $s = \epsilon$, hvor ϵ er det minste tillatte inkrement (for eksempel 1 krone i Norge eller 1 cent i USA).

La oss tenke oss at vi vil undersøke om adferden i ultimatumeksperimentet varierer med x. Mer presist ønsker vi å undersøke «high-stakes-hypotesen»: om F får et stort nok pengebeløp til fordeling, senker F andelen som tilbys M. Dette fordi F forventer en høyere akseptrate fra M når det ligger mer på bordet. I eksperimentet benytter vi et moderat beløp $\underline{x}$ (kontroll) og et skikkelig høyt beløp $\overline{x}$ (behandling). Til eksperimentet rekrutterer vi to grupper subjekter: én blir kontrollgruppe, den andre behandlingsgruppe. Vi deler så hver av disse gruppene inn i to nye grupper, én gruppe med F-spillere og én gruppe med M-spillere. Deretter danner vi par av F-spillere og M-spillere. Så gjennomføres ultimatumspillet i hvert av parene. Vi kan nå reformulere hypotesen vår: den gjennomsnittlige andelen som F tilbyr M, synker når vi går fra kontrollgruppe til behandlingsgruppe.

Behandlingseffekt

La $i = 1, 2, \ldots n$ være F-spillere. En behandlingseffekt (eller, om du vil, kausaleffekt) defineres som forskjellen mellom *potensielle utfall* for et individ.[1] I vårt tilfelle er det to potensielle utfall for hver *i*: $s_i(\overline{x})$ og $s_i(\underline{x})$. Mens $s_i(\overline{x})$ angir den andelen *i* tilbyr når hun er i behandlingsgruppen (høyt beløp til fordeling), så angir $s_i(\underline{x})$ den andelen *i* tilbyr når hun er i kontrollgruppen (moderat beløp til fordeling). Effekten på sjenerøsiteten av å ha et høyt pengebeløp til fordeling, *behandlingseffekten*, kan skrives:

$$\tau_i = s_i(\overline{x}) - s_i(\underline{x})$$

Men nå kan det se ut som om vi er i «big trouble»: Behandlingseffekten kan ikke uten videre bestemmes fra data; enhver F-spiller er enten i kontrollgruppen eller i behandlingsgruppen, aldri i begge grupper samtidig.

Med utgangspunkt i behandlingseffekten til det enkelte subjekt kan vi definere en *gjennomsnittlig* behandlingseffekt:[2]

$$ATE = E\left[s_i(\overline{x})\right] - E\left[s_i(\underline{x})\right]$$

ATE tillater de potensielle bidragene til ulike individer å variere, og beregner en gjennomsnittlig behandlingseffekt basert på de potensielle bidragene.

---

1 Notasjonen og definisjonene i dette avsnittet bygger på Rubin (1974).

2 Vi følger den engelskspråklige litteraturen og skriver ATE («average treatment effect»). $E[\cdot]$ står for forventning («expectation»).

ATE er et sentralt begrep; det forteller oss hvilken forventet (altså gjennomsnittlig) effekt det har når alle F går fra kontrollgruppe til behandlingsgruppe. I vårt eksempel er dette den gjennomsnittlige effekten på sjenerøsiteten til F hvis vi lar *hver og en* av dem gå fra et moderat til et høyt beløp å fordele.

Problemet er fortsatt at F-spillerne enten er i kontrollgruppen eller i behandlingsgruppen, aldri i begge grupper. Kausalitet ser altså ut til å forutsette kunnskap om kontrafaktiske utfall. Hvis vi kunne observere hver enkelt F både i kontrollgruppen og i behandlingsgruppen, kunne vi beregnet kausaleffekten for hver F, for så å beregne *ATE*.[3]

Gruppetilordning

La oss fordele subjektene i eksperimentet tilfeldig på rollene som F og M, deretter fordele F og M tilfeldig på kontroll- og behandlingsgruppe, for til sist å skape tilfeldige par av F og M i begge grupper.

Hvorvidt *i* er i kontroll- eller behandlingsgruppen er nå en stokastisk variabel. Vi benevner denne variablen D_i. $D_i = 1$ hvis *i* tilfeldigvis havnet i behandlingsgruppen, og $D_i = 0$ hvis *i* tilfeldigvis havnet i kontrollgruppen.

Med tilfeldig tildeling til de to gruppene (tilfeldig trekk av utfall på D_i) er forventet potensiell sjenerøsitet fra F-spillere i behandlingsgruppen (den betingede forventningen) lik forventet potensiell sjenerøsitet fra alle F-spillere (den ubetingede forventningen). Eller, med symboler: $E\left[s_i\left(\overline{x}\right) \mid D_i = 1\right] = E\left[s_i\left(\overline{x}\right)\right]$. Hvorfor? Fordi F-spillere i behandlingsgruppen er trukket tilfeldig fra gruppen av alle F-spillere. Siden F-spillere i kontrollgruppen også er et tilfeldig trukket utvalg fra gruppen av alle F spillere, må vi ha at $E\left[s_i\left(\overline{x}\right) \mid D_i = 0\right] = E\left[s_i\left(\overline{x}\right)\right]$. Et parallelt resonnement gir $E\left[s_i\left(\underline{x}\right) \mid D_i = 0\right] = E\left[s_i\left(\underline{x}\right)\right]$ og $E\left[s_i(\underline{x}) \mid D_i = 1\right] = E\left[s_i(\underline{x})\right]$.

Ved å tilordne subjekter tilfeldig til grupper og roller sørger vi for at all heterogenitet over subjekter, både den som er observerbar og den vi ikke kan observere, er tilfeldig fordelt i gruppene. Forventningsmessig kan det da ikke være noen forskjell mellom F spillere trukket fra kontroll- og behandlingsgruppen. Vi har ikke grunnlag for å forvente noe annet enn at en F spiller i behandlingsgruppen ville ha opptrådt som en F spiller i kontrollgruppen, dersom vi hadde plassert ham i kontrollgruppen. Tilsvarende må vi forvente at en F spiller i kontrollgruppen ville ha opptrådt som en F spiller i behandlingsgruppen, dersom vi hadde plassert henne i behandlingsgruppen.

Aktiv bruk av tilfeldighet har gitt oss kontroll over de kontrafaktiske utfallene! Vi har nå det som trengs for å estimere ATE basert på *observert*

3 Vi skal se på et design som tilstreber dette om litt. Foreløpig holder vi fast ved at forslagsstillere er i én og bare én gruppe.

adferd. Hvis vi lar m være antall F-spillere i behandlingsgruppen, og $n - m$ være antall F-spillere i kontrollgruppen, har vi at:

$$ATE = E\left[s_i\left(\overline{x}\right) \mid D_i = 1\right] - E\left[s_i\left(\underline{x}\right) \mid D_i = 0\right]$$
$$= \tfrac{1}{m}\sum_{i=1}^{m} s_i - \tfrac{1}{n-m}\sum_{i=m+1}^{n} s_i$$

Tilfeldig tilordning tillater oss altså å estimere den gjennomsnittlige behandlingseffekten fra data, ved å ta differansen mellom gjennomsnittlig sjenerøsitet i behandlings- og kontrollgruppen. Dette er svært kraftfullt. Hvis vi for eksempel observerer $ATE < 0$, og forskjellen er statistisk signifikant, så kan vi si at høy belønning gjennomsnittlig sett *for årsaker* mindre sjenerøsitet i eksperimentet. Vi har ikke bare funnet en korrelasjon, men har påvist en årsak til effekten vi observerer. Nøkkelen er tilfeldig tilordning.

Uten tilfeldig tilordning må behandlingseffekter estimeres på andre måter. Disse måtene er beheftet med problemer og er langt mer usikre. Økonometrikere som analyserer ikke-eksperimentelle data har derfor begynt å spørre seg hvordan en gitt problemstilling ville blitt belyst i et ideelt eksperimentelt design. Når de har tenkt dette ut, spør de hvilken metode som ligger tettest mulig opp til det idelle eksperiment (for eksempel Angrist & Pischke 2009).

Seleksjon

Hva om vi ikke tilordner F-spillerne tilfeldig til kontroll- og behandlingsgruppe, men lar dem velge gruppe selv, ut fra et førstemann-til-mølla-prinsipp? Vi kan tenke oss at de mest grådige av dem ville bruke en del energi på å trenge seg frem i køen til behandlingsgruppen, der de kan tjene mer penger.

Den gjennomsnittlige observerte endringen i sjenerøsitet mellom behandlingsgruppen og kontrollgruppen er:[4]

$$E\left[s_i\left(\overline{x}\right) \mid D_i = 1\right] - E\left[s_i\left(\underline{x}\right) \mid D_i = 0\right] =$$
$$E\left[s_i\left(\overline{x}\right) \mid D_i = 1\right] - E\left[s_i\left(\underline{x}\right) \mid D_i = 1\right] +$$
$$E\left[s_i\left(\underline{x}\right) \mid D_i = 1\right] - E\left[s_i\left(\underline{x}\right) \mid D_i = 0\right]$$

Differansen i rekke to er den gjennomsnittlige effekten av at subjekter i behandlingsgruppen gis behandling (=forventet forskjell i potensielle utfall blant de som fikk behandling). Differansen i den siste rekken er en seleksjonseffekt (=forventet forskjell i potensielle utfall av *ikke* å bli behandlet for sub-

4 Alt som er gjort i ligningen, er å legge til og trekke fra $E\left(s_i\left(\underline{x}\right) \mid D = 1\right)$ etter likhetstegnet.

jekter i behandlingsgruppen og i kontrollgruppen). Med tilfeldig tilordning til kontroll -og behandlingsgruppe har vi sett at denne differansen er null.[5]

Når subjektene selv har selektert til gruppene, gjelder ikke dette lenger. Subjektene i gruppene kan være systematisk forskjellige med hensyn på de potensielle utfallene. Med utgangspunkt i køargumentet vårt mistenker vi spesielt at $E\left[s_i\left(\underline{x}\right) \mid D_i = 1\right] < E\left[s_i\left(\underline{x}\right) \mid D_i = 0\right]$, og at vi derfor har en negativ seleksjonseffekt. Det vil si: subjekter i behandlingsgruppen ville ha vært mindre sjenerøse enn subjekter i kontrollgruppen selv om de også hadde hatt en liten sum å fordele. Estimatet vil i så fall overdrive størrelsen på den negative effekten av å øke beløpet (x) som skal fordeles. Når tilordning til grupper ikke er tilfeldig, kan vi ikke estimere behandlingseffekten ved å ta differansen mellom faktisk sjenerøsitet i behandlings- og kontrollgruppen.

Innen og mellom

Hva om vi deler gruppen av subjekter i to, og lar alle subjekter spille ultimatumspillet to ganger etter hverandre? Vi gjør det slik: halvparten av subjektene er F-spillere, resten er M-spillere. Halvparten av F-spillerne er i kontrollgruppen i første spill, den andre halvparten er i behandlingsgruppen. I annet spill bytter F-spillerne gruppe; de som var i kontrollgruppen i første spill, er i behandlingsgruppen i annet spill, og de som var i behandlingsgruppen i første spill er i kontrollgruppen i annet spill. Med dette opplegget vil vi observere hva den enkelte F-spiller tilbyr både når vedkommende er i kontrollgruppen og når vedkommende er i behandlingsgruppen. Løser ikke dette problemet vårt?

Nei! Hvorfor ikke? Fordi F-spillere ikke kan betraktes som «de samme spillerne» i spill én og i spill to. Erfaringene fra første spill kan virke inn på adferden i spill to. Vi må derfor randomisere over rekkefølgen som F-spillerne er i kontroll- og behandlingsgruppe.

Men vi kan ikke bare tenke på rekkefølgen til F-spillerne. Vi må også lage et design for rekkefølgen til M-spillerne. Det sikreste er å uttømme alle muligheter ved å sørge for å samle data fra de fire mulige kombinasjonene:

1. F-spillere som starter med $\overline{x}$ og i spill to møter M-spillere som startet med $\overline{x}$
2. F-spillere som starter med $\overline{x}$ og i spill to møter M-spillere som startet med $\underline{x}$
3. F-spillere som starter med $\underline{x}$ og i spill to møter M-spillere som startet med $\underline{x}$
4. F-spillere som starter med $\underline{x}$ og i spill to møter M-spillere som startet med $\overline{x}$

For å kunne estimere behandlingseffekter, må vi randomisere over disse fire behandlingene. Hvis vi ikke gjør det, kan den observerte effekten skyldes

5 Siden $E\left[s_i\left(\underline{x}\right) \mid D_i = 0\right] = E\left[s_i\left(\underline{x}\right)\right]$ og $E\left[s_i\left(\underline{x}\right) \mid D_i = 1\right] = E\left[s_i\left(\underline{x}\right)\right]$ så er $E\left[s_i\left(\underline{x}\right) \mid D_i = 1\right] - E\left[s_i\left(\underline{x}\right) \mid D_i = 0\right] = 0$

rekkefølgen behandlingen introduseres i. «High-stakes-hypotesen» sier nå at s_i øker fra første til annet spill i 1. og 2., og reduseres fra første til annet spill i 3. og 4. Vi må også sjekke at det ikke er signifikante rekkefølgeforskjeller mellom 1. og 2., og heller ikke mellom 3. og 4.

Når effekter estimeres ved å utnytte endringen i adferd for samme subjekt over tid, kalles dette et «innen-subjekt-design». Når effekter estimeres ved å utnytte endringen i adferd for ulike subjekter på samme tidspunkt kalles dette et «mellom-subjekt-design».

I designet over testes high-stakes-hypotesen i et innen-subjekt-design, mens rekkefølgeeffektene testes i et mellom-subjekt-design. Konklusjonen er at vi ikke kan droppe randomisering selv om vi undersøker samme subjekt i ulike behandlinger på ulike tidspunkter.

Kontroll

Laboratoriet åpner for utpreget kontroll med omgivelsene beslutningene tas i. Når man er interessert i å teste en presis modell, er dette en vesentlig fordel. I dette avsnittet presenterer vi først en modell. Deretter diskuterer vi hvordan incitamenter, institusjoner og informasjonsstruktur fra denne modellen kan gjenskapes i laboratoriet. Det er dette vi forstår med kontroll.

Etableringsspill

Et interessant spørsmål er hva som kan tenkes å drive over- og underetableringer i et markded. Dette er et omfattende spørsmål som kan presiseres i mange retninger. Vi benytter en stilisert modell til å analysere *én enkelt* side av slike etableringer, nemlig det *koordineringsproblemet* homogene bedrifter står overfor når markedet er for lite til å gi profitt ved overetablering (Selten & Güth 1982).[6]

I modellen tar $i = (1, 2, \ldots, n)$ spillere («bedriftene») hver for seg en beslutning om hvorvidt de skal etablere seg i markedet eller ikke. Spillet spilles bare én gang, og spillerne tar sine beslutninger uten å kjenne de andre spillernes beslutning (de velger «simultant»). Det er ingen anledning til å kommunisere før beslutning fattes.

En spiller som ikke går inn i markedet, får en sikker gevinst. Gevinsten til en spiller som går inn, er en lineært synkende funksjon av antallet andre spillere som går inn. Spilleregler og belønninger er åpent kjent. Mer presist ser belønningsfunksjonen slik ut:

6 Vi kan tenke på modellen som en «redusert form-modell». Vi antar at bedriftene etter etablering spiller sine likevektsstrategier i et oligopolspill mot andre bedrifter som også har etablert seg. Oligopolspillet modelleres ikke direkte, men profitten synker med antallet konkurrenter som er etablert i markedet, slik tilfellet også vil være i en oligopollikevekt.

$$v_i(e_i, e_{-i}) = \begin{cases} a & hvis \quad e_i = 0 \\ b + r(k-m) & hvis \quad e_i = 1 \end{cases}$$

Valget til spiller *i* symboliseres med $e_i = \{0,1\}$, hvor $e_i = 1$ betyr at spiller *i* går inn. Belønningen til spiller *i* er en funksjon av hans eget valg, e_i, men også av de andre spillernes valg, e_{-i}. Det totale antallet spillere som går inn, er gitt ved ved $m = \sum_{i=1}^{n} e_i$. Markedets kapasitet er gitt ved *k*, mens avkastningen av en etablering er gitt ved *r*. I modellen er *k*, *r*, *a* og *b* konstanter.

Hva er Nash-likevekten(e) i spillet? En Nash likevekt krever at ingen spiller har grunn til å angre sitt strategivalg, gitt de andre spillernes valg av strategi. Vi sier at spillernes strategivalg må være bestesvar mot hverandre. En strategi er en forskrift som forteller spilleren hva hun skal gjøre i alle valgsituasjoner hun kan møte i spillet. Siden vi ser på et spill med simultane valg, kan ingen spiller betinge sine handlinger av andre spilleres adferd. En strategi vil da være identisk med en handling. Hver spiller har altså to strategier: gå inn eller stå utenfor.

La oss se på tilfellet der *m* spillere har gått inn, mens $n - m$ spillere står utenfor. En spiller som har gått inn har ikke grunn til å angre dette hvis $b + r(k-m) \geq a$, altså hvis $m \leq \frac{rk+b-a}{r}$. En spiller som står utenfor har ikke grunn til å angre dette hvis $b + r(k-(m+1)) \leq a$, altså hvis $m \geq \frac{r(k-1)+b-a}{r}$. For at spillernes valg skal være bestesvar mot hverandre må altså (hvor * symboliserer «i likevekt»):

$$\frac{rk+b-a}{r} \geq m^* \geq \frac{r(k-1)+b-a}{r}$$

La oss sette $a = b = 1$ og $r = 2$. Med disse parametrene unngår vi heltallsproblemer og Nash-betingelsen reduserer til $k \geq m^* \geq (k-1)$. Hvis for eksempel $k = 5$, så finnes det en Nash-likevekt hvor 5 spillere går inn, og en annen Nash-likevekt hvor 4 spillere går inn.

Det er lett å innse at det må være slik. Hvis jeg står utenfor og 5 andre spillere er inne, får jeg 0 hvis jeg også går inn $(1 + 2(5-6))$. Hvis jeg ikke går inn, får jeg 1. Hvis jeg er en av dem som gikk inn, får jeg 1 hvis jeg blir $(1 + 2(5-5))$ og 1 hvis jeg går ut. Når 5 spillere har gått inn, kan hverken de som er inne eller de som står utenfor profitere på å endre sitt valg.Ved å regne litt, kan du selv forsikre deg om at tilsvarende gjelder for $k - 1 = 4$, men ikke for andre verdier på *k* enn 4 og 5.

Det finnes altså to Nash-likevekter i rene strategier. Imidlertid finnes det mer enn én måte å spille dem på. Likevekten hvor m^* spillere går inn sier ingen ting om *hvilke* m^* spillere dette er. Det er $\binom{n}{k}$ ulike måter å spille

likevekten med $m^* = k$.[7] I tillegg er det $\binom{n}{k-1}$ måter å spille likevekten med $m^* = k - 1$.

Det finnes i tillegg også en likevekt i blandede strategier. I denne likevekten velger alle spillerne å gå inn i markedet med sannsynlighet $p_i^* = \frac{k-1}{n-1}$.[8]

Hvis vi bare betrakter de rene strategiene kan vi regne ut at med $n = 10$, så er det 462 forskjellige måter å spille i likevekt på, mens det med $n = 20$ er 20 349 forskjellige måter å spille i likevekt på. Det er tilsynelatende en uoverkommelig oppgave for spillere (som ikke engang tillates å kommunisere) å koordinere på en av de mange likevektene i spillet. Spillteorien kan heller ikke by på noen akseptert modell for hvilken av måtene å spille på som vil bli valgt.

Men hvor vanlig er det egentlig at koordineringsproblemet leder til over- og underetableringer? Og hvilke prinsipper ser faktiske mennesker ut til å benytte for å lykkes i sin koordinering?

Feltdata

Om vi søker svar ved å konfrontere modellresonnementet med innsamlede data om faktiske markeder (ikke-eksperimentelle feltdata), støter vi på flere utfordringer. Disse er blant annet knyttet til måleproblemer, endogenitet og seleksjon på uobserverbare variable.

Modellen sier at en gitt kapasitet i markedet forklarer etableringsadferden. I markeder utenfor laboratoriet kan vi imidlertid ikke utelukke at kausaliteten går motsatt vei, ved at bedriftene påvirker markedets kapasitet. Etableringshindringer skyldes ofte myndighetsreguleringer. Disse kan påvirkes gjennom lobbyfremstøt. Etablerte bedrifter kan også se seg tjent med å prise lavt, investere for å komme nye patenter i forkjøpet, og skape kundelojalitet gjennom markedsføring, i den hensikt å gjøre nyetableringer uprofitable.

---

7 $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ er telleregelen (binomialkoeffisienten) som forteller oss hvor mange koalisjoner med k medlemmer som kan dannes i en gruppe på tilsammen n medlemmer.

8 Likevekt i blandede strategier behandles i kapittel 2. Argumentet for likevekten i blandede strategier i etableringsspillet er som følger: Når hver av de andre spillerne går inn i markedet med sannsynlighet p_{-i} er min forventning til m, gitt at også jeg går inn, lik $E_i\left(m \mid e_i = 1\right) = 1 + \left(n-1\right)p_{-i}$. For at jeg skal være indifferent mellom å gå inn eller stå utenfor, må $b + r\left(k - E_i\left(m \mid e_i = 1\right)\right) = a$. Med våre parametre, og innsatt for $E_i\left(m \mid e_i = 1\right)$, blir indifferensbetingelsen $1 + 2\left(k - 1 + \left(n-1\right)p_{-i}\right) = 1$. En enkel mellomregning gir $p_{-i}^* = \frac{k-1}{n-1}$. Siden jeg er indifferent mellom handlingsvalgene mine når alle andre spillere randomiserer over sine handlingsvalg med p_{-i}^*, er ethvert svar et bestesvar. Hvis jeg i likhet med alle andre spillere velger å sette $p_i^* = \frac{k-1}{n-1}$, så er dette en likevekt – når alle spillerne går inn i markedet med sannsynlighet p_i^*, kan ingen spiller profitere på å gjøre noe annet, gitt det de andre spillerne gjør.

I markeder utenfor laboratoriet er det umulig å identifisere alle de bedriftene som *kan* ha vurdert å etablere seg. Det vi observerer er etablerte bedrifter. Dette reiser betydelige måleproblemer.

Mange ting varierer over faktiske bedrifter. For eksempel bedriftenes etableringskostnader, deres produksjonsteknologi, deres tilgang til produkt-, kapital- og arbeidsmarkeder, måten de har organisert seg på, deres bedriftskultur, og deres renomme i markedet. Bedrifter vil stå overfor mye usikkerhet knyttet til alle sider ved sine potensielle konkurrenter. Alle disse forholdene kan påvirke bedriftenes etableringsbeslutninger. Vi må derfor gå ut fra at de bedriftene som etablerer seg, ikke utgjør et representativt utvalg av bedrifter. Etablerte bedrifter er en selektert gruppe, og seleksjonen er (i alle fall delvis) basert på uobserverbare forhold. Dette skaper, som vi har sett, utfordringer med hensyn til å estimere kausaleffekter.

Etableringsbeslutninger tas ofte i en sekvens. Det kan derfor være at etableringene i markedet vi studerer beskrives bedre av andre modeller enn den vi har valgt å se på. Mens laboratorieeksperimenter vanligvis sier noe om adferd innenfor gitte institusjoner, er det mer sjeldent å se eksperimentelle studier av hvilke institusjoner som faktisk benyttes i ulike samhandlingssituasjoner utenfor laboratoriet.[9]

Modellen vi ser på bygger på restriktive forutsetninger, og den fokuserer på én type markedsinstitusjon. Dette er bevisst. Hensikten er å rendyrke det koordineringsproblemet som oppstår når etableringsbeslutninger tas simultant av homogene bedrifter, for å vurdere hvor stort *dette* problemet er. Det er denne modellen vi ønsker å teste, ikke en annen modell som bygger på andre markedsinstitusjoner. Dette er normalvitenskap. Vi bryter opp et problemkompleks i mindre problemer som lar seg takle. La oss se nærmere på hvordan modellen kan testes på en kontrollert måte i et laboratorieeksperiment.

Redelighet

En forskjell mellom laboratorieeksperimenter i økonomi og psykologi er at mens økonomer ikke forleder subjektene sine, så gjør psykologer det rutinemessig.

La oss tenke oss at vi gjennomfører et eksperiment for å teste etableringsspillet, og oppgir at belønningen ved å gå inn er $1+2(5-m)$ og at $n = 20$, men at subjektene ikke tror på dette. Da vil det ikke lenger gi god mening å undersøke koordinering i etableringsspillet under den forutsetning at alle subjekter kjenner belønningsstrukturen.

Et ekstremt eksempel på bruk av forledning er autoritetseksperimentet til Yale-psykologen Stanley Milgram (1963). I eksperimentet satt subjektet –

9 Se kapittel 4 om endogene strafferegimer for et eksempel på at også valg av institusjoner lar seg studere i laboratoriet.

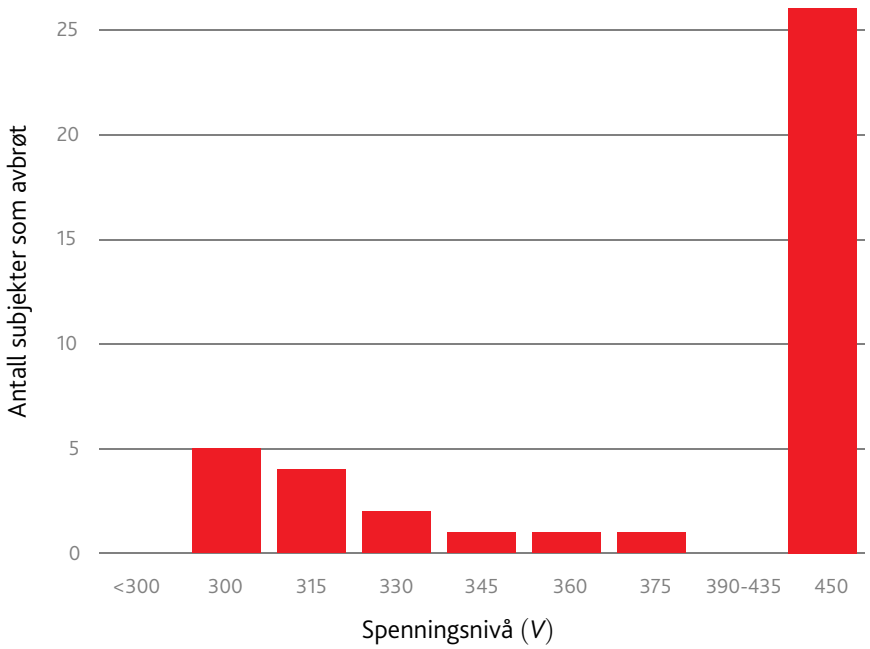
benevnt «lærer» – i et rom sammen med eksperimentets administrator (kledd i en teknikeruniform og instruert til å opptre med streng mine og ikke vise følelser). Gjennom et vindu inn til et annet rom kunne de se en innleid skuespiller – benevnt «elev» – som var koplet opp til elektroder. Høytalere og mikrofoner i de to rommene sørget for lydforbindelse.

I eksperimentet ble subjektene forledet til å tro at rollene som «lærer» og «elev» var tilfeldig fordelt, og at subjektet hadde trukket rollen som «lærer». Subjektet («læreren») ble – uriktig – fortalt at han bidro til forskning på sammenhengen mellom straff og læring.

Subjektet («læreren») skulle lese opp en oppgave, som skuespilleren («eleven») så svarte på. Ved feil svar skulle subjektet administrere et elektrisk støt til skuespilleren på en skala fra 45 til 450 volt. De øverste trinnene var merket «fare – svært alvorlig støt». For å troverdiggjøre narreriet, ble subjektet selv utsatt for et ekte, men mildt støt på 45 volt før eksperimentet startet.

Skuespilleren var imidlertid ikke koplet opp til aktiv spenning, og fikk derfor ikke noe «støt» når subjektet brukte skalaen. Skuespilleren kunne se hvor mye «støt» subjektet administrerte, og hadde et skript for å uttrykke økende smerte, verbalt og med kroppsspråk, ved stigende «støt».

Hvis subjektet kviet seg for å administrere «støt» til skuespilleren, hadde administrator et skript med eskalerende oppfordringer om å gjøre det: «vær så snill å fortsette»; «eksperimentet krever at du fortsetter, vær så snill å



Figur 1.1 Autoritetseksperiment (*Basert på Milgram 1963*)

fortsette»; «det er helt avgjørende at du fortsetter»; «du har ikke noe annet valg, du må fortsette». Dersom subjektet fortsatt nektet å gi «støt» etter den fjerde oppfordringen, ble sesjonen avsluttet. Den virkelige hensikten med eksperimentet var å teste ut menneskers autoritetstro.

Resultatene fra eksperimentet er vist i figur 1.1. Ingen av de 40 subjektene som deltok i eksperimentet, nektet å administrere «støt» til skuespilleren. Tvert i mot viste samtlige subjekter vilje til å administrere «støt» som var mellom 36 og 105 prosent kraftigere enn det du får ved å stikke to strikkpinner inn i et norsk støpsel. Hele 65 prosent av subjektene ble med hele veien til 450 volt.

Eksperimentet er ekstremt og har i ettertid blitt kritisert for å være uetisk, særlig fordi det utsatte subjektene for betydelig stress. At subjektene ble forledet (med hensyn til forskningens formål, randomiseringen over roller, og realiteten i «støtene») er ikke noe eksperimentelle psykologer ser ut til å ha så store problemer med.

Normen i eksperimentell økonomi er radikalt forskjellig. Forledning *skal* ikke benyttes. Eksperimentelle arbeider som har brukt forledning, har liten eller ingen sjanse til å bli publisert i økonomiske fagtidsskrifter.

Hvorfor? Begrunnelsen er todelt. For det første leder forledning til tap av kontroll. Subjekter som har blitt forledet, eller som har hørt at forskerne i det aktuelle laboratoriet har brukt forledning, vil forvente å bli forledet. Med slike forventninger tapes kontrollen med subjektenes oppfatninger. Subjektene må danne seg sine egne oppfatninger om de reelle parameterne i eksperimentet, og disse oppfatningene vil variere.

Den andre delen av begrunnelsen er at forledning forurenser subjektpoolen (de subjektene som er villig til å delta i fremtidige eksperimenter). For forskere må poolen av subjekter betraktes som et fellesgode (se kapittel 4). En enkeltforsker kan tenke at forledning er veien til interessante resultater (autoritetseksperimentet til Milgram ga for eksempel *svært* interessante resultater). Hvis noen forskere benytter forledning, vil subjektene forvente å bli forledet, og dette vil ødelegge fremtidige forskningsresultater for *alle* forskere. For å disiplinere forskere til ikke å bruke forledning, er det et kraftfullt virkemiddel å nekte arbeider som har brukt forledning publisering. Et laboratorium for eksperimentell økonomi er avhengig av å etablere og vedlikeholde et unntaksløst rykte for å ikke forlede.

Incentiver

I psykologien benyttes vanligvis ikke monetære belønninger for å knytte incentiver til handlingsvalg i eksperimenter. Dette til forskjell fra eksperimentell økonomi, hvor monetære belønninger så å si alltid benyttes.

Hvorfor? Incentiver benyttes for å oppnå kontroll med subjektetenes preferanser. Subjektene som dukker opp i laboratoriet, har nyttefunksjoner

som forskeren ikke observerer. Vi tenker oss at disse funksjonene har tre argumenter: det subjektet tjener gjennom sine handlinger i eksperimentet, Δx, penger subjektet har fra før, x_0, og alt annet som får subjektet til å handle, **z**. Vi skriver den uobserverte nyttefunksjonen som $U(x_0, \Delta x, \mathrm{z})$. For at subjektet skal ha incentiver som samsvarer med modellen forskeren undersøker («induced preferences»), må det stilles to krav (Smith 1982):

1. *Monotoni*: U er strengt voksende i x for alle kombinasjoner av x_0 og **z**
2. *Prominens*: Endringer i subjektets nytte fra eksperimentet kommer hovedsakelig fra Δx, og innflytelsen til **z** er neglisjerbar.

Hvis vi lager et eksperiment basert på etableringsspillet, er det lett å forestille seg at studenter som deltar kan tilfredstille disse kravene. For studenter er x_0 som regel lav relativt til de Δx som typisk benyttes i eksperimenter. Det er også tvilsomt om etableringsspillet inviterer til å handle med utgangspunkt i moralnormer eller sosiale normer, så av denne grunn er innflytelsen til **z** er rimeligvis liten.

Hvis vi tror Δx er for liten, kan den økes, enten ved å øke nivået på belønningene, hvis eksperimentet gjennomføres i et rikt land, eller ved å gjennomføre eksperimentet i et land hvor en krone har større kjøpekraft (se kapittel 3). I noen typer eksperimenter kan det være grunn til å tro at **z** har større innflytelse (se kapittel 3 og 4 for eksempler).

Å søke forklaringer på eksperimentelle resultater i **z** er en viktig del av det feltet som kalles «adferdsøkonomi».

Legg merke til at incentiver oppstår når subjektene betales for *handlingene* de tar i eksperimentet. Eksperimentet incentiveres altså ikke ved at det betales en fast sum for å møte opp, noe som ikke er uvanlig i eksperimenter i psykologi.[10]

10 I praksis incentiveres eksperimenter omtrent på denne måten: vi vil lage et eksperiment med etableringsspillet og bruke parameterne $a = b = 1$, $k = 5$, $r = 2$ og $n = 20$. Vi annonserer at subjektene vil tjene en eksperimentell valuta ECU («Experimental Currency Unit») i eksperimentet, og at denne vil bli vekslet inn etter en fast og forhåndsannonsert kurs etter at eksperimentet er sluttført. Hvis subjektene spiller likevekten med $m^* = 4$, vil dette koste 28 ECU per runde. Hvis de spiller likevekten med $m^* = 5$, vil det koste 20 ECU per runde. Antar vi at likevektene er like sannsynlige, vil forventet kostnad i likevekt være 24 ECU per runde. Vi vet hvor lang tid en periode tar (fordi vi setter en maksimal tid som subjektene tillates å bruke per steg i programvaren). For et gitt antall perioder (=repetisjoner av grunnspillet) vet vi hva den forventede utbetalingen i ECU er hvis subjektene er i likevekt. Vi kan da sette vekslekursen slik at studentene forventer å tjene godt hvis de spiller i likevekt. I laboratoriet til BI benyttes for tiden (2014) en forventet utbetaling i likevekt