

Samarbeid via inspeksjon

Jon Reiersen

Tønsberg: Høgskolen i Vestfold, 2004

Notat 4/2004

Notat 4/2004 Høgskolen i Vestfold
Copyright: Høgskolen i Vestfold/Jon Reiersen
ISSN: 0808-131x

Samarbeid via inspeksjon

JONREIERSEN¹

Sammendrag

Artikkelen analyserer hvordan samarbeid kan etableres i sosiale dilemma situasjoner. Det fokuseres på det som kan tolkes som anonyme markedsrelasjoner der partene møtes tilfeldig og bare en gang - og hvor det ikke finnes eksterne håndhevelsesmekanismer. Strategier endres over tid via kulturell evolusjon der handlingsmønstre som oppfattes som suksessfulle av medlemmene i samfunnet sprer seg over tid. Det vises at en norm om samarbeid ikke er nok til å sikre at samarbeid blir etablert. En norm om samarbeid koplet med muligheten til å undersøke hvilken "type" en samhandler med kan imidlertid sikre at samarbeid blir etablert. Størrelsen på kostnaden knyttet til å undersøke hvem en samhandler med er med på å bestemme graden av samarbeid i samfunnet. Dersom denne kostnaden er liten vil andelen som samarbeider bli stor, men det er vanskelig å få etablert en stabil situasjon hvor alle samarbeider. Dette resultatet holder for ulike spesifikasjoner av den underliggende sosiale interaksjonen.

JEL klassifisering: A12, A13, C72, Z13

Emneord: Sosiale dilemma, normer, samarbeid, kulturell evolusjon

¹Takk til Kalle Moene for kommentarer til et tidligere utkast av artikkelen.

“[M]odern life leads individuals to face the frequent anonymous, nonrepeated interactions that are characteristic of modern societies with advanced trade, communication, and transportation technologies.” (Gintis m.fl., s. 168, 2003).

1 Innledning

Samarbeid muliggjør arbeidsdeling og spesialisering som er en viktig kilde til vekst og velstand. I mange sammenhenger stilles vi imidlertid overfor sterke materielle incentiver til å bryte formelle eller uformelle avtaler om samarbeid. Ved ikke å samarbeide er det mulig å sikre seg selv en gevinst på bekostning av andre, men dersom alle forfølger en slik strategi blir det ingen gevinst å dele. Thomas Hobbes var en av de første som formulerte dette sosiale dilemmaet eksplisitt. Han betraktet naturtilstanden nettopp som en situasjon der alle drives av å sikre sine egne egoistiske interesser på bekostning av andre. Han argumenterte for at dersom denne egeninteressen ikke temmes vil vi ledes inn i en kollektiv katastrofe hvor livet er “(...) *solitary, poor, nasty, brutish, and short.*” (Hobbes 1651, s. 100). For å komme ut av naturtilstanden trenger vi en ytre disiplinerte organ som har kraft til å temme våre opportunistiske tendenser og stimulere til samarbeid for fellesskapets beste. En sterk statsmakt med velfungerende institusjoner som straffer brudd på inngåtte avtaler kan være et slikt organ, i følge Hobbes. Men selv i moderne siviliserte samfunn med velutviklet lovverk, rettsinstanser og politi eksisterer det mange muligheter til å fremme egne materielle interesser på bekostning av andre ved å bryte avtaler om samarbeid. I mange sammenhenger kan det oppstå uforutsette forhold som det er vanskelig å spesifisere på forhånd i en kontrakt som skal styre transaksjonen mellom to parter. De fleste avtaler eller kontrakter er derfor ufullstendige. Dersom uforutsette forhold inntreffer gir det rom for at minst en av partene kan utnytte dette til egen fordel. Dette til tross, et framtrødende trekk ved de fleste samfunn er at selv i situasjoner med sterke materielle incentiver til å bryte formelle eller uformelle avtaler om samarbeid, så blir slik avtaler i stor grad overholdt (Richerson m.fl., 2003).

Et forhold som kan forklare dette er at mye handel og annen sosial samhandling foregår innenfor rammen av langvarige og stabile relasjoner. Langvarige relasjoner kan stimulere til samarbeid fordi den kortsiktige gevinsten ved å “svike” motparten i dag ikke er stor nok til å kompensere for det langsiktige tapet ved at et slikt svik vil lede motparten til å avvike sitt samarbeid i framtiden. Samarbeid er med andre ord forenlig med rasjonelle

aktørers langsiktige egeninteresse.² Men som sitatet innledningsvis indikerer så eksisterer det mange situasjoner hvor betingelsene for å etablere samarbeid gjennom gjentatt interaksjon ikke er til stede. Handel i markeder med mange tilbydere og etterspørrere er et eksempel. Hva er det som motiverer den enkelte til å holde avtaler (samarbeide) innenfor slike anonyme relasjoner, og hva er det som sikrer at samarbeid vedlikeholdes over tid under fravær av gjentatt interaksjon mellom partene? Formålet med denne artikkelen er å analysere dette spørsmålet.

Mekanismen jeg vil fokusere på er muligheten for *inspeksjon*. Med inspeksjon mener jeg at den enkelte kan skaffe seg informasjon om en partner en i utgangspunktet ikke kjenner ved å undersøke hvordan denne har oppført seg i tidligere relasjoner. Fordelen med slik informasjon er at partene kan bruke denne som beslutningsgrunnlag ved valg av egne handlinger. Dersom en for eksempel ønsker å samarbeide, så er det mindre risikofylt å velge samarbeid dersom en har informasjon om at en potensiell partner er til å stole på. På den måten kan samarbeid vokse fram. Men informasjon om partnerens “type” kan også brukes opportunistisk for å fremme egne materielle interesser. Ved å velge ‘ikke samarbeid’ når en vet at partneren samarbeider, vil en sikre seg selv en gevinst på bekostning av partneren.

Muligheten for å etablere samarbeid i sosiale dilemma situasjoner basert på inspeksjon er diskutert og analysert av flere (Hamilton 1963; Frank 1988, 1989; Skyrms 1994; Stark 1999, 2004; Güth og Kliemt 1998, 2000; Bowles og Gintis 2004). Disse arbeidene bygger imidlertid på svært ulike antakelser om den underliggende sosiale interaksjonen som studeres. En målsetting med denne artikkelen er derfor å forsøke og integrere den eksisterende litteraturen innenfor et felles rammeverk. Hensikten er spesielt å klargjøre hvilke underliggende antakelser som driver fram resultatene som litteraturen referert til ovenfor har produsert. Som vi skal se nedenfor så er disse resultatene svært følsomme overfor endringer i antakelser knyttet til beslutningsstruktur og de involverte partenes motivasjon. Konklusjonen som følger fra analysen nedenfor er likevel entydig. Muligheten for inspeksjon vil under plausible betingelser føre til at samarbeid kan etableres som et stabilt utfall i en sosial dilemma situasjon.

Et mindre avklart spørsmål i litteraturen referert til ovenfor er hva som motiverer den enkelte til å samarbeide i en sosial dilemma situasjon. En slik situasjon er jo nettopp kjennetegnet ved at det er best for den enkelte å

²Dette er analysert inngående både innenfor klassisk spillteori (Friedman 1971; Fudenberg og Maskin 1986) og evolusjonær spillteori (Trivers 1971; Axelrod og Hamilton 1981).

velge 'ikke samarbeid' selv om det for de involverte er best om alle samarbeider. Både hverdagserfaringer og eksperimenter gjennomført av psykologer, økonomer og andre samfunnsvitere peker imidlertid i retning av at folk flest er innstilt på å samarbeide i sosiale dilemma situasjoner - også i situasjoner der det ikke er rom for gjentatt interaksjon (Sober og Wilson, 1998; Fehr og Gächter, 2002; Camerer, 2003; Richerson m.fl. 2003 og Henrich m.fl. 2004). Sosiologer har lenge pekt på *normer* for å forklare dette. Normer om samarbeid holder våre opportunistiske tendenser i sjakk, og leder oss til å handle i tråd med fellesskapets interesser.³ Økonomer har imidlertid vært lite villige til å innføre normer for å forklare økonomisk eller sosial adferd. En vanlig innvending er at en enhver handling i prinsippet vil kunne forklares ved å peke på en eller annen norm. Å innføre normer for å forklare en handling eller et fenomen er derfor nesten som å anta svaret.⁴ Nedenfor vil jeg likevel innføre normer for å forklare hvordan samarbeid kan etableres i en sosial dilemma situasjon. Jeg vil imidlertid forsøke å *forklare* eksistensen av disse normene som et resultat av en pågående sosial interaksjon. Dette gjøres ved å anta et samfunn bestående av to typer aktører. Den ene typen, som jeg kaller samarbeidere, er motivert av en internalisert norm om samarbeid. For disse typene er det en verdi i seg selv å samarbeide. Den andre typen, som jeg kaller opportunister, er ikke opptatt av å handle i tråd med normer om samarbeid. Opportunistene vil bare velge å samarbeide dersom dette fremmer deres egne materielle interesser. Disse to aktørtypene utgjør medlemmene av samfunnet som vi skal betrakte. Innenfor dette samfunnet møtes medlemmene for å spille ulike varianter av et fangenes dilemma

³Se Horne (2001) for en gjennomgang av sentrale bidrag til den sosiologiske litteraturen som behandler begrepet 'normer'. Borgstede (2002) analyserer normer i sosiale dilemma situasjoner mer spesielt, mens Elster (1989) og Torsvik (2003) gir en mer generell diskusjon av sammenhengen mellom normer, rasjonalitet og egennyttig handling.

⁴Det interessante spørsmålet er derfor ikke først og fremst hva som er konsekvensen av ulike typer normer, men heller hvorfor ulike normer eksisterer. Som Coleman (1990) kommenterer: "*Much sociological theory takes norms as given and proceeds to examine individual behavior or the behavior of social systems when norms exist. Yet to do this without raising at some point the question of why and how norms come into existence is to forsake the more important sociological problem in order to to address the less important.*" (Coleman, 1990 s. 244). I den grad normer blir forsøkt forklart benyttes gjerne såkalte funksjonelle forklaringer der en norm (eller andre fenomener) blir forklart ved hjelp av den gunstige effekten normen (eller fenomenet) har for en gruppe eller for et samfunn. Problemet med en slik forklaring er at et fenomen forklares ved hjelp av et annet fenomen som inntreffer senere i tid. Det faktum at en norm om samarbeid er nyttig fordi den leder folk til å samarbeide er interessant, men altså ikke tilstrekkelig som en forklaring på at normen eksisterer. Se Elster (1983) for en mer utfyllende kritikk av funksjonelle forklaringer.

spill. Partene møtes kun en gang og de velger handling i henhold til sine subjektive preferanser. Disse preferansene endres imidlertid over tid via en sosial evolusjonær prosess. Det antas at hver enkelt forsøker å tilpasse sine handlinger over tid til den handlingen som gir størst utbytte i den sosiale interaksjonen som studeres. Implikasjonene av dette er at “overlevelsessevnen” til en norm om samarbeid er bestemt av utbyttet som følger av den handlingen som normen tilskriver. Dersom handlinger som bryter med normen om samarbeid generer et høyere utbytte enn det å holde seg til normen, så vil det være en tendens til at normen om samarbeid brytes over tid også av de som i utgangspunktet var “bærere” av den. Det motsatte vil skje dersom handlinger som er i tråd med normen om samarbeid genererer et høyere utbytte enn handlinger som bryter med normen. Normer om samarbeid vil da spre seg over tid blant aktørene i samfunnet.⁵

Artikkelen er disponert på følgende måte. I avsnitt 2 og 3 presiserer jeg nærmere den sosiale interaksjonen som studeres og hvordan preferanser endres over tid. I avsnitt 4 til 8 studerer jeg ulike varianter av den underliggende sosiale interaksjonen der det overordnede spørsmålet er om en norm om samarbeid, koblet med mulighet for inspeksjon, kan føre til at samarbeid etableres som en stabil situasjon. Noen avsluttende kommentarer følger i avsnitt 9.

2 Den sosiale interaksjonen

Anta et samfunn bestående av et stort antall aktører som møtes parvis. Det er tilfeldig hvem en møter og hvert møte er kun en engangsforeteelse. Spillerne velger derfor handling uavhengig av forventninger om framtidig interaksjon. Mulige handlinger for den enkelte aktør (som heretter kalles spiller) og konsekvenser av disse handlingene er angitt i spillmatrisen under.

	Samarbeid	Ikke samarbeid
Samarbeid	β , β	$-\alpha$, α
Ikke samarbeid	α , $-\alpha$	0 , 0

Anta at $\alpha > \beta > 0$, det vil si at spillerne er involvert i et fangenes dilemma spill. Spillet kan for eksempel tolkes som en markedstransaksjon der den ene spilleren er selger og den andre spilleren en kjøper. Dersom begge

⁵Se også Basu (1995), Sethi (1996) og Bendor og Swistak (2001) som benytter en lignende tilnærming for å forklare eksistensen av normer.

velger ‘samarbeid’ så betyr det at begge parter oppfyller sine forpliktelser. Selgeren leverer varen han har forpliktet seg til å levere, mens kjøperen betaler den avtalte prisen. Begge mottar et utbytte på β . Dersom begge velger ‘ikke samarbeid’ så betyr det at selgeren ikke leverer varen og kjøperen heller ikke betaler den avtalte prisen. Det vil si at et ikke finner sted noen transaksjon, og begge får null i utbytte. Dersom selgeren ikke leverer varen (velger ‘ikke samarbeid’), mens kjøperen betaler den avtalte prisen (velger ‘samarbeid’), så får selgeren utbyttet α , mens kjøperen får det lavest mulige utbyttet $-\alpha$. Dette fordi selgeren sparer kostnaden ved å framstille varen, samtidig som han får betalt for den. Kjøperen kommer da dårlig ut siden han betaler for en vare han ikke mottar. Dersom det motsatte tilfellet inntreffer, altså at selgeren leverer varen men ikke får betaling fra kjøperen, så er det selgeren som får $-\alpha$ mens kjøperen får det høyeste utbyttet α .

Dette gir oss det velkjente resultatet fra fangenes dilemma spill. Siden ‘ikke samarbeid’ gir det høyeste utbyttet for hver av spillerne uansett hvilken handling den andre spilleren velger, blir utfallet at ingen av partene samarbeider. Begge får dermed null i utbytte som er lavere enn utbyttet de hadde fått dersom begge hadde samarbeidet (som ville gitt β). Individuell rasjonalitet leder altså til kollektiv urasjonalitet.

Innenfor rammen av dette spillet ønsker jeg nå å undersøke nærmere hvordan muligheten til å undersøke hvem en samhandler med kan føre til at samarbeid vokser fram. For å kunne gjøre dette er det imidlertid nødvendig å si noe mer om hva som kan motivere den enkelte til å velge samarbeid i situasjonen ovenfor, og dernest hvordan denne motivasjonen kan tenkes å bli endret over tid. Jeg starter med det siste først.

3 Kulturell evolusjon

Inspirert av evolusjonær spillteori skal jeg anta at handlingsmønstre overføres fra et individ til et annet gjennom imitasjon eller en annen form for sosial læring. Spesielt vil jeg anta at handlingsmønstre som oppfattes som suksessfulle av medlemmene i samfunnet vil spre seg over tid.⁶ Med suksessfull vil jeg videre mene at vedkommende handling gir et høyt utbytte i den sosiale interaksjonen (spillet) som medlemmene av samfunnet er involvert i.

En slik evolusjonær prosess kan presiseres mer formelt på følgende måte. Anta et samfunn bestående av et stort antall individer som kan deles opp i

⁶Se Boyd og Richerson (1985), og Richerson, Boyd og Henrich (2003) og Bowles (2003, kap. 2) for en mer utfyllende diskusjon og analyse av en slik “kulturell” evolusjonær prosess.

to typer, *samarbeidere* (S -typer) og *opportunister* (O -typer). Hvilken type en er bestemmer hvilken handling en velger i spillet en er involvert i. La $p \in (0, 1)$ være andelen samarbeidere i samfunnet der \dot{p} angir veksten i andelen samarbeidere over tid. Dersom spillerne møtes parvis, og det er tilfeldig hvem en møter, så kan forventet utbytte til de to typene skrives

$$U(S) = pm(S, S) + (1 - p)m(S, O) \quad (1)$$

$$U(O) = pm(O, S) + (1 - p)m(O, O) \quad (2)$$

der $m(i, j)$ er utbyttet til en i -type når han spiller mot en j -type. Relasjon (1) leses på følgende måte: Med sannsynlighet p møter en S -type en annen S -type noe som gir utbyttet $m(S, S)$, mens med sannsynlighet $(1 - p)$ møter den samme S -Typen en O -type noe som gir utbyttet $m(S, O)$. Relasjon (2) har en tilsvarende tolkning.

På ethvert tidspunkt vil veksten i andelen samarbeidere i samfunnet være positiv eller negativ avhengig av om forventet utbytte til samarbeiderne er større eller mindre enn forventet utbytte til opportunistene.⁷ La $D(p)$ angi differansen mellom forventet utbytte til de to typene når andelen samarbeidere i samfunnet er p , det vil si

$$D(p) = U(S; p) - U(O; p) \quad (3)$$

Vi har da at $\dot{p} \gtrless 0$ alt etter som $D(p) \gtrless 0$, det vil si at andelen samarbeidere vil vokse eller avta over tid avhengig av om de har et høyere eller lavere forventet utbytte enn opportunistene.

Vi kan tolke dynamikken skissert ovenfor på følgende måte: Samfunnet starter ut med en gitt fordeling av samarbeidere og opportunister som møtes tilfeldig. Hver spiller velger handling i henhold til hvilken type de er noe som igjen tilskriver hver av spillerne et bestemt utbytte. Etter at den sosiale interaksjonen har funnet sted vurderer en andel av spillerne om de skal endre sitt valg av handling. Gitt at en spiller vurderer å endre handling, så er denne beslutningen basert på en sammenligning av eget utbytte med utbyttet som den alternative handlingen genererer. Hvis den alternative handlingen genererer et høyere utbytte enn det spilleren får selv, så vil spilleren bytte til den alternative handlingen. Hvis den alternative handlingen generer et lavere utbytte så vil spilleren holde på sitt opprinnelige valg av handling.

⁷Weibull (1995) kaller dette antakelsen om *payoff monotonicity*. For å studere payoff monoton dynamikk trenger vi med andre ord kun å sammenligne forventet utbytte til de to typene.

Likevekt, $\dot{p} = 0$, oppnås når $D(p) = 0$, det vil si at forventet utbytte til de to typene er den samme. En slik likevekt (p^*) kan være en av følgende tre typer. Dersom $p^* = 0$ har vi et samfunn kun bestående av opportunister, dersom $p^* = 1$ har vi et samfunn kun bestående av samarbeidere og dersom $p^* \in (0, 1)$ har vi et blandet samfunn bestående av både samarbeidere og opportunister. En likevekt kan videre være stabil eller ustabil. Dersom p^* tilfredsstiller $D(p^*) = 0$, så er dette en stabil likevekt dersom en liten bevegelse bort fra p^* leder til en endring i andelen samarbeidere tilbake til p^* . Stabilitet krever altså at dersom andelen samarbeidere øker, så øker det forventede utbyttet for opportunistene relativt til samarbeiderne, noe som igjen leder til en økning i andelen opportunister.

4 Normer om samarbeid

Hva kan forklare at noen vil velge å samarbeide innenfor et fangenes dilemma spill der den sosiale interaksjonen foregår kun en gang med den samme personen? Som nevnt ovenfor har sosiologer vært opptatt av å vise hvordan normer kan begrense folks tilbøyelighet til å handle egennyttig og opportunistisk. La oss derfor i spillet ovenfor introdusere ideen om at noen av aktørene er motivert av en norm om samarbeid som sier at: “*Det er galt å ikke samarbeide*”. La oss kalle de som er “bærere” av denne normen for *samarbeidere* (*S*-typer). La oss videre anta normen om samarbeid er fullstendig internalisert hos de som bærer den, det vil si at de vil ta hensyn til den selv om handlinger som representerer et brudd med normen ikke blir oppdaget og sanksjonert av andre. Motivet for å handle i tråd med normen om samarbeid er ønsket om å unngå følelsen av skyld, skam og tap av selvrespekt.⁸ La $\tau > 0$ angi den subjektive sosio-psykologiske kostnaden ved å bryte normen om samarbeid. Når to samarbeidere møtes vil dermed spillmatrisen bli seende ut som illustrert nedenfor.

	Samarbeid	Ikke samarbeid
Samarbeid	β , β	$-\alpha$, $\alpha - \tau$
Ikke samarbeid	$\alpha - \tau$, $-\alpha$	$0 - \tau$, $0 - \tau$

Vi ser at hvis $\tau > \alpha$ så blir ‘samarbeid’ en dominerende strategi for spillerne. Dersom $\tau > \alpha$ så vil altså to samarbeidere når de møtes velge

⁸Bowles og Gintis (2001) gir en interessant diskusjon av forhold som kan forklare hvorfor vi er utstyrt med evnen til å internalisere normer, og hvilken rolle følelser som skyld og skam spiller for økonomisk handling. Se også Gintis (2003).

‘samarbeid’ og få β i utbytte. De som ikke har internalisert normen om samarbeid har $\tau = 0$, og vil dermed ha ‘ikke samarbeid’ som dominerende strategi. Jeg kaller disse for *opportunist* (*O*-typer). Når to opportuniste møtes vil altså begge velge ‘ikke samarbeid’ og få 0 i utbytte.

Hvordan vil andelen opportuniste og samarbeidere utvikle seg over tid når vi antar at spillernes handlinger utvikler seg i tråd med den evolusjonære prosessen beskrevet i avsnitt 3, og spillerne *ikke* har mulighet til å undersøke hvem de møter?⁹ Når andelen samarbeidere er p blir forventet utbytte til samarbeiderne $U(S) = p\beta + (1 - p)(-\alpha)$, mens opportunistene får $U(O) = p\alpha - (1 - p)0$. Setter vi dette inn i (3) finner vi at $D(p) = p\beta - \alpha$. Siden $\alpha > \beta$ vil $D(p) < 0$ for alle $p \in (0, 1)$ og $\dot{p} < 0$. Det vil altså si at dersom vi starter ut med et samfunn bestående av både samarbeidere og opportuniste, så vil andelen samarbeidere avta over tid inntil $p = 0$. Den eneste stabile evolusjonære likevekten er at ingen samarbeider.

Forklaringen på dette resultatet er at når samarbeiderne er drevet av en betingelsesløs norm om samarbeid, så vil de samarbeide uansett hvem de møter. Dette høster opportunistene en fordel av siden de da får det høyeste utbytte α hver gang de møter en samarbeider. Opportunistene utnytter altså samarbeidernes udiskriminerende samarbeid og får et høyere forventet utbytte. Dette vil lede til en situasjon der samarbeiderne gradvis endrer handling til ‘ikke samarbeid’. Denne prosessen vil pågå helt til det ikke er noen igjen som samarbeider ($p = 0$).

5 Samarbeidere, opportuniste og inspektører

Konklusjonen ovenfor, at samarbeid ikke er en levedyktig strategi, må imidlertid modifiseres dersom vi innfører en ny type i samfunnet, nemlig *inspektører* (*I*-typer). Inspektøren er en type som i likhet med samarbeiderne er innstilt på å samarbeide, men som først vil undersøke nærmere hvem han møter før han faktisk velger å samarbeide. I det følgende skal jeg også anta at slik inspeksjon med sikkerhet avslører hvem en står overfor. Det vil si at ved å undersøke vil en inspektør vite om han spiller mot en samarbeider, en opportunist eller en annen inspektør. Men siden inspeksjon som regel tar tid og krefter, så skal jeg også anta at en slik handling påfører inspektørene

⁹Den evolusjonære prosessen beskrevet ovenfor innebærer at selv om samarbeiderne bryr seg om normen om samarbeid, så er “overlevelsesnivået” til normen bestemt av utbyttet som følger av den handlingen normen tilskriver. Dette innebærer at dersom handlinger som bryter med normen om samarbeid generer et høyere utbytte, så er det en tendens til at normen om samarbeids brytes også av de som i utgangspunktet var “bærere” av den.

en kostnad K .¹⁰

La som før andelen samarbeidere i samfunnet være p , mens andelen inspektører er q . Andelen opportunister blir da $(1 - p - q)$. Vi holder også fast på antakelsen om at det er tilfeldig hvem en møter for å spille spillet definert ovenfor bortsett fra at inspektørene nå kan avsløre hvem de spiller mot, og dermed velge strategi deretter.

I utgangspunktet er det grunn til å tro at introduksjonen av inspektører i samfunnet er gunstig for samarbeiderne siden samarbeidere får flere å samarbeide med. Hvorvidt introduksjonen av inspektører gjør det mulig for samarbeid å “overleve” kan vi undersøke ved å sammenligne forventet utbytte til de tre typene, som vil være

$$U(I) = p\beta + q\beta + (1 - p - q)0 - K = (p + q)\beta - K \quad (4)$$

$$U(S) = p\beta + q\beta + (1 - p - q)(-\alpha) = (p + q)\beta - (1 - p - q)\alpha \quad (5)$$

$$U(O) = p\alpha + q0 + (1 - p - q)0 = p\alpha \quad (6)$$

Fra avsnitt 3 vet vi at likevekt krever at forventet utbytte til de tre typene er den samme. (4) og (5) gir da

$$1 - p - q = \frac{K}{\alpha} \quad (7)$$

hvor $(1 - p - q) \in (0, 1)$ dersom $K < \alpha$. (5) og (6), sammen med (7) gir videre

$$q = \frac{1}{\alpha} \left[(\alpha - \beta) \left(1 - \frac{K}{\alpha} \right) + K \right] \quad (8)$$

der $q > 0$ når $K < \alpha$. Ved å sette (8) inn i (7) får vi så at andelen samarbeidere blir

$$p = \frac{1}{\alpha} \left[\beta - (\alpha + \beta) \frac{K}{\alpha} \right] \quad (9)$$

¹⁰Modellen nedenfor er en variant av modellene studert nærmere av Bowles og Gintis (2004) og Stark (2004). Min framstilling skiller seg fra Bowles og Gintis ved at de antar at det ikke medfører kostnader for inspektørene å undersøke, samt at en undersøkelse fra inspektørens side bare avslører de andre spillernes type med sannsynlighet p . Stark antar at inspektørene ikke har en interaksjon før de møter en samarbeider eller en annen inspektør. Det betyr at dersom inspektørene møter en opportunist første gang spillet spilles, så vil de vente til neste “runde” av spillet for å søke etter en de kan samarbeide med der. I framstillingen nedenfor antar jeg at *alle* spillerne har en interaksjon i hver runde spillet spilles, men at inspektørene tilpasser sin strategi til hvem de møter. Resultatene fra modellene til Bowles og Gintis (2004) og Stark (2004) er imidlertid kvalitativt sett de samme som de som følger nedenfor.

Fra (9) følger da at $p > 0$ dersom $\beta > (\alpha + \beta) K/\alpha$, eller dersom

$$K < \frac{\alpha\beta}{\beta + \alpha} \equiv K_{\max} \quad (10)$$

Siden (10) også impliserer $K < \alpha$, så er (10) en nødvendig og tilstrekkelig betingelse for en likevekt der alle tre typene er representert. Fra (8) og (9) finner vi videre at

$$\frac{dq}{dK} = \frac{\beta}{\alpha^2} > 0 \quad \text{og} \quad \frac{dp}{dK} = -\frac{\alpha + \beta}{\alpha^2} < 0 \quad (11)$$

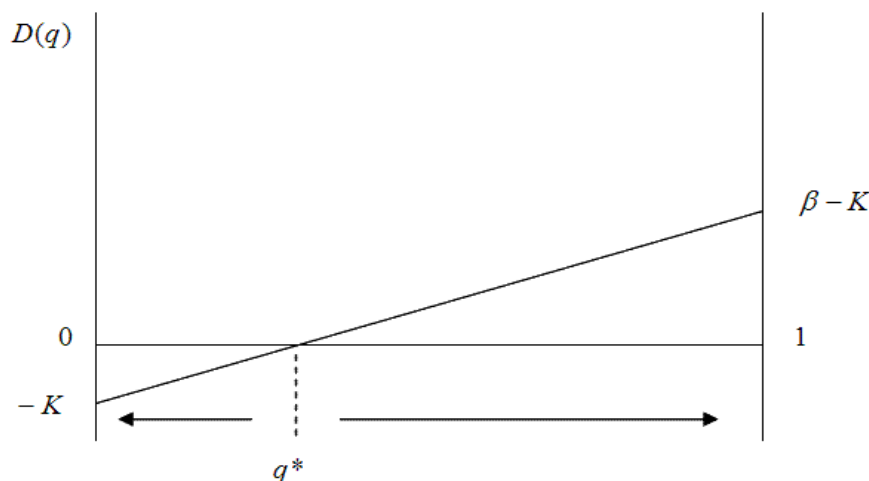
som betyr at når kostnaden ved å undersøke øker så vil andelen inspektører øke, og andelen samarbeidere reduseres. Fra (7) ser vi også at andelen opportunister øker (den totale andelen av inspektører og samarbeidere går ned). Det vil igjen si at graden av samarbeid i samfunnet synker.

Et talleksempel kan illustrere dette nærmere. Anta at $\alpha = 4$ og $\beta = 2$. Fra (10) får vi da at $K_{\max} = 1.33$. Kostnaden ved å undersøke kan altså ikke være større enn 1.33 for å ha et samfunn der samarbeid er en levedyktig strategi. La kostnaden ved å undersøke (K) være 0.5. Fra (7)-(9) får vi da en likevekt som inneholder $p = 31\%$ samarbeidere, $q = 56\%$ inspektører og $1 - p - q = 13\%$ opportunister. Dersom vi setter $K = 1$ vil andelen opportunister øke til 25%, mens andelen inspektører øker til 63%. Andelen samarbeidere vil dermed gå ned til 12%.

Konklusjonen er altså at dersom kostnadene ved å undersøke ikke er for stor ($K < \alpha\beta/(\alpha + \beta)$) så er det mulig å etablere et samfunn der alle de tre typene beskrevet ovenfor er representert. Jo større K er desto større er andelen opportunister som igjen betyr mindre grad av samarbeid.

Analysen i avsnitt 4 viste videre vist at det ikke er mulig med en likevekt bestående av bare opportunister og samarbeidere (selv med $K = 0$). Det er heller ikke er mulig med en likevekt bestående av kun samarbeidere og inspektører siden samarbeiderne alltid vil gjøre det bedre når det ikke er behov for å beskytte seg mot opportunistene. (De blir spart for kostnaden K .) Det er derimot mulig med en likevekt bestående av opportunister og inspektører. For å studere denne muligheten nærmere la q være andelen inspektører i samfunnet, mens $(1 - q)$ er andelen opportunister. Disse har samme type adferd som beskrevet ovenfor (opportunistene velger alltid 'ikke samarbeid' mens inspektørene undersøker og samarbeider med andre inspektører, mens de velger 'ikke samarbeid' dersom de møter en opportunist). Forventet utbytte for inspektørene blir dermed $q\beta + (1 - q)0 - K$. Opportunistene blir møtt med 'ikke samarbeid' uansett om de møter en av samme type eller

en inspektør, og får dermed null i utbytte. Benytter vi dette i (3) får vi at $D(q) = q\beta - K$. Likevekt oppnås når $D(q) = 0$, som gir $q^* = K/\beta$ der $q^* \in (0, 1)$ hvis $K < \beta$. Men denne indre likevekten er ustabil siden $D'(q) = \beta > 0$. En liten økning i q bort fra q^* øker det forventede utbyttet til inspektørene relativt til opportunistene, noe som igjen leder til en ytterligere økning i andelen inspektører. Situasjonen er illustrert i figur 1.



Figur 1 viser at det eksisterer tre evolusjonære likevekter, nemlig 0, q^* og 1, hvorav den første og siste er stabile. Den indre likevekten q^* definerer et kritisk “vippepunkt” i den forstand at straks andelen inspektører er større enn q^* så vil andelen inspektører vokse inntil $q = 1$. En reduksjon i kostnaden ved å undersøke flytter q^* mot venstre. Det kreves dermed en mindre andel inspektører initialt for å starte en utvikling i retning av den stabile likevekten $q = 1$.^{11,12}

¹¹Dersom vi setter $\beta = 2$ og $K = 1.5$ får vi $q^* = 0.75$. Reduserer vi K slik at $K = 0.5$ får vi $q^* = 0.25$.

¹²Skyrms (1994) betrakter en variant av situasjonen ovenfor der han antar at det ikke er forbundet med kostnader ved å undersøke hvem en møter ($K = 0$), og at spillet består av to “runder”. I første runde møtes spillerne tilfeldig. Dersom en samarbeider møter en annen samarbeider så samarbeider de. Hvis en samarbeider møter en opportunist, så finner det ikke sted noen interaksjon. Det samme gjelder dersom en opportunist møter en annen opportunist (opportunistene vil gjerne unngå hverandre like mye som samarbeiderne vil unngå dem). Alle medlemmene av populasjonen som ikke finner en partner i første runde møtes tilfeldig igjen i andre runde. Alle gir nå opp å undersøke og har isteden en interaksjon med den de tilfeldigvis møter. Utstyrt med den samme evolusjonære prosessen beskrevet i avsnitt 3 ovenfor, viser Skyrms at også denne situasjonen leder til en ustabil indre likevekt á la den som er illustrert i figur 3.

Modellen illustrerer dermed at den samme underliggende sosiale interaksjonen kan lede til to svært forskjellige stabile situasjoner. Hvor samfunnet ender opp avhenger av hvor det starter (for gitt payoff struktur og for en gitt kostnad ved å undersøke). Dersom den initiale andelen inspektører er liten (mindre enn q^*) så vil samfunnet bevege seg i en retning der ingen samarbeider. Dersom den initiale andelen inspektører er stor (større enn q^*) så vil samfunnet bevege seg i en retning der alle samarbeider. Begge situasjonene er mulig, det vil si at den evolusjonære prosessen er *historie-betinget*.¹³

6 Betinget samarbeid

I modellen ovenfor er spillernes adferd helt mekanisk i den forstand at de velger handling avhengig av hvilken type de er og (på kort sikt) uavhengig av hvordan omgivelsene ser ut (hvilke handlinger de andre velger). Frank (1988; 1989) bryter med denne antakelsen ved å studere en situasjon der beslutningen om å inspisere eller ikke er basert på rasjonelle overveielser.

I modellen til Frank er det to typer spillere, *opportunistene* og *samarbeiderne*. Opportunistene samarbeider aldri (som ovenfor), mens inspektørene ønsker å samarbeide med andre som også samarbeider.¹⁴ Vi kan tolke dette som at inspektørene er drevet av en *betinget* norm om samarbeid som sier at: “*Det er galt å ikke samarbeide med en som samarbeider*”.¹⁵ Vi kan framstille en slik betinget norm om samarbeid i fangenes dilemma spillet ovenfor på følgende måte. La som før $\tau > 0$ angi den subjektive sosio-psykologiske kostnaden ved å bryte normen om samarbeid, der det kun er samarbeiderne

¹³En slik historie-betinget evolusjonær prosess står sentralt i Marx sin historieoppfatning, noe som blant annet følgende kommentar i *The Eighteenth Brumaire of Louis Bonaparte* (1852) uttrykker: “*Men make history, but they do not make it just as they please; they do not make it under circumstances chosen by themselves but under circumstances (...) given and transmitted from the past. The tradition of all the dead generations weights like a nightmare on the brain of the living.*” (Marx, 1852(1963) s. 15). Modellen ovenfor, og variantene av den samme modellen som analyseres nedenfor, kan betraktes som et eksempel på hvordan en slik prosess kan modelleres mer formelt. Rollen samfunnets historie spiller for utviklingen av normer og institusjoner, og for hvordan et samfunnet ser ut “i dag”, er også studert nærmere av North (1990), Putnam (1993), Platteau (2000), Rothstein (2000) samt Hechter og Opp (2001). Se også Bowles 2004 (kap. 1 og 12).

¹⁴Modellen som presenteres i Stark (1999) er identisk med modellen til Frank (1988; 1989) uten at leseren gjøres oppmerksom på dette. Guttman (2000) analyserer en modell som er beslektet med modellen til Frank, men der det åpnes for at opportunistene kan velge ‘samarbeid’ dersom det fremmer deres interesser. Opportunistene er med andre ord ikke “programmert” til å velge ‘ikke samarbeid’ slik som i framstillingen til Frank. Muligheten for at også opportunistene kan velge ‘samarbeid’ blir studert nærmere i avsnitt 8.

¹⁵Se Opp (2001) for en diskusjon av forskjellen mellom en *betinget* og en *ubetinget* norm.

som er bærere av en slik norm. For opportunistene er $\tau = 0$. Når to inspektører med en betinget norm om samarbeid møtes vil dermed spillmatrisen endres til

	Samarbeid	Ikke samarbeid
Samarbeid	β , β	$-\alpha$, $\alpha - \tau$
Ikke samarbeid	$\alpha - \tau$, $-\alpha$	0 , 0

Vi ser at dersom $\tau > \alpha - \beta$ så vil begge spillerne ønske å samarbeide bare dersom den andre samarbeider. Det opprinnelige spillet endres fra å være et fangenes dilemma spill til å bli et koordineringsspill. Dette spillet har to Nash-likevekter. Enten velger begge spillerne ‘samarbeid’ og får (β, β) , eller begge velger ‘ikke samarbeid’ og får $(0, 0)$.¹⁶ I det følgende skal jeg anta at når to samarbeidere møtes så vil de samarbeide.¹⁷ Siden opportunistene har $\tau = 0$ vil de spille fangenes dilemma spillet når de møtes. Begge velger ‘ikke samarbeid’ og får 0 i utbytte.

Som nevnt ønsker samarbeiderne inspektørene å samarbeide med andre samarbeidere. Problemet for samarbeiderne er imidlertid at de ikke vet hvilken type de møter, de er bare i stand til å observere andelen samarbeidere i samfunnet. Men som i avsnitt 5 skal vi anta at samarbeiderne kan skaffe seg informasjon om hvilken type de møter ved å inspisere. En slik inspeksjon vil avsløre med sikkerhet partnerens sanne type, men dette koster K . Mens inspektørene i modellen analysert i avsnitt 5 undersøkte uansett størrelsen på K og p , vil samarbeiderne nå basere sin beslutning om å undersøke eller ikke på hva dette koster av seg. Det vil si at de bare vil undersøke dersom det forventede utbyttet ved en slik strategi er større enn utbyttet ved å ikke undersøke.

Andelen opportunister og samarbeidere utvikler seg fortsatt i henhold til den evolusjonære prosessen beskrevet i avsnitt 3, men der det også tas hensyn til at samarbeiderne kjenner til hvordan størrelsen på p utvikler seg

¹⁶Det eksisterer også en likevekt i blandede strategier der begge spillerne velger å samarbeide med sannsynlighet $\alpha / (\beta + \tau)$.

¹⁷Denne antakelsen kan begrunnes mer formelt ved å introdusere begrepet *risiko dominant* likevekt (se for eksempel Young (1998)). Anta at begge spillerne er usikre på hvilken strategi den andre vil velge. Dersom begge tror at den andre velger samarbeid med sannsynlighet 0.5, så blir forventet utbetaling ved å samarbeide $(\beta - \alpha) / 2$, mens forventet utbetaling ved ikke å samarbeide blir $(\alpha - \tau - 0) / 2$. Dersom spillerne er opptatt av å maksimere forventet nytte så vil de velge ‘samarbeid’ dersom $(\beta - \alpha) / 2 > (\alpha - \tau - 0) / 2$, eller dersom $\tau > 2\alpha - \beta$. Dette er det samme som å si at samarbeid er risikodominant. Dersom $\tau > 2\alpha - \beta$ er oppfylt så vil altså to samarbeidere velge ‘samarbeid’ når de møtes.

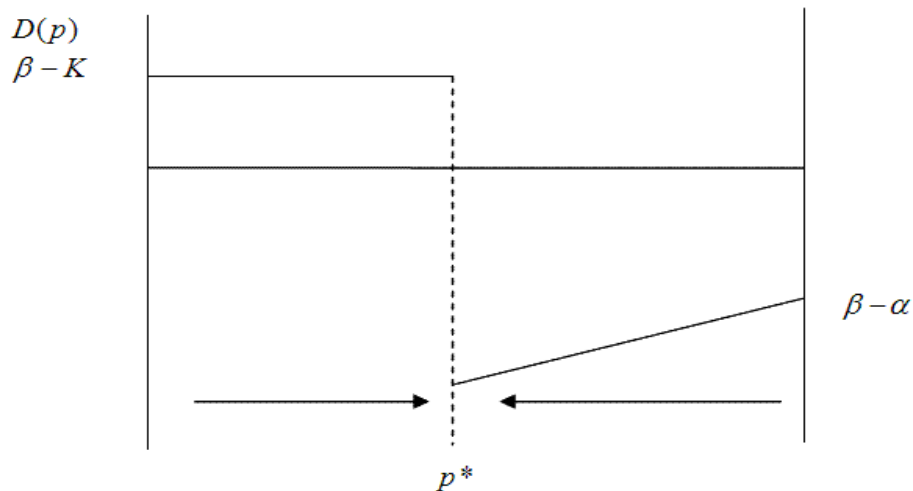
over tid. Dynamikken i spillet blir da som følger. Samfunnet starter ut med en gitt fordeling av opportunister og samarbeidere. Disse møtes parvis der det er tilfeldig hvem en møter. Etter at den sosiale interaksjonen har funnet sted endrer sammensetningen av populasjonen seg i henhold til det relative utbyttet til de to typene i samfunnet. Før neste runde spilles skaffer samarbeiderne seg så informasjon om andelen samarbeidere i samfunnet. Spillerne møtes så tilfeldig igjen og velger handling i tråd med sine subjektive preferanser, etter at samarbeiderne også har hatt mulighet til å skaffe seg informasjon om hvilke type partneren er.

Spørsmålet er da under hvilke betingelser samarbeiderne vil skaffe seg informasjon om hvilke type partneren er. Ved å undersøke vil samarbeiderne finne fram til andre samarbeidere og få $\beta - K$ med sikkerhet.¹⁸ Opportunistene blir overlatt til å spille med seg selv og får 0 i utbytte. Samarbeiderne vil dermed gjøre det bedre enn opportunistene dersom $K < \beta$. Samtidig har samarbeiderne også mulighet til å velge samarbeid uten å inspisere, noe som vil spare dem for kostnaden ved å undersøke. Dette vil gi samarbeiderne et forventet utbytte på $p\beta + (1 - p)(-\alpha)$. Samarbeiderne vil derfor velge å undersøke dersom $\beta - K > p\beta - (1 - p)\alpha$. Løser vi med hensyn på p finner vi at samarbeiderne vil undersøke dersom

$$p < 1 - \frac{K}{\alpha + \beta} = p^* \quad (12)$$

hvor $p^* \in (0, 1)$ dersom $K < \alpha + \beta$. Siden denne betingelsen også inneholder $K < \beta$, så er kravet for at samarbeiderne undersøker at $K < \beta$. Dersom denne ulikheten er oppfylt og spillernes handlinger utvikler seg i tråd med den evolusjonære prosessen beskrevet i avsnitt 3, så kan utviklingen i p over tid framstilles ved hjelp av figur 2.

¹⁸Her er det altså antatt at samarbeiderne aldri vil møte en opportunist dersom samarbeideren undersøker.



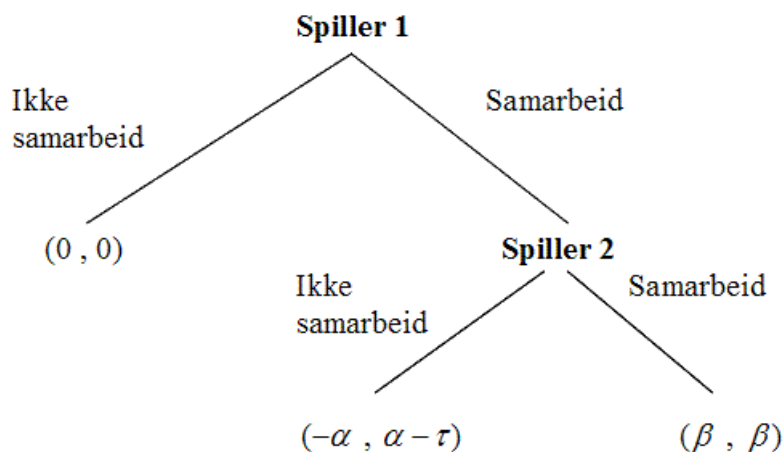
Dersom $p < p^*$ vil samarbeiderne undersøke hvem de møter og samarbeider kun dersom de møter en annen samarbeider. Samarbeidernes forventede utbytte blir $U(S) = \beta - K$, mens opportunistene får $U(O) = 0$. Benytter vi dette i (3) får vi at $D(p) = \beta - K > 0$ og $\dot{p} > 0$ siden $K < \beta$. Andelen samarbeidere vil dermed øke inntil $p = p^*$. Når $p > p^*$ er andelen som samarbeider i samfunnet så stor at samarbeiderne velger å samarbeide betingelsesløst framfor å undersøke. Dette gir dem $U(S) = p\beta + (1 - p)(-\alpha)$, mens opportunistene får $U(O) = p\alpha + (1 - p)0$. Innsatt i (3) gir dette $D(p) = p\beta - \alpha < 0$ og $\dot{p} < 0$ siden $\alpha > \beta$. Når $p > p^*$ utnytter opportunistene samarbeidernes udiskriminerende samarbeid, og får et høyere forventet utbytte. Andelen samarbeidere vil dermed avta over tid inntil $p = p^*$. p^* er derfor en stabil indre likevekt. Samfunnet vil derfor aldri nå en tilstand der samarbeiderne “tar over” hele samfunnet. Samarbeiderne vil heller leve side om side med en andel $(1 - p^*)$ ikke samarbeidende opportunister.

Et talleksempel kan igjen illustrere dette nærmere. Anta at $\alpha = 4$, $\beta = 2$ og $K = 1$. Fra (12) får vi da en likevekt som inneholder $p^* = 83\%$ samarbeidere og 17% opportunister. Dersom vi øker K til 1.5 reduseres andelen samarbeidere til $p^* = 75\%$, mens en reduksjon i K til 0.5 øker andelen samarbeidere til $p^* = 92\%$. Andelen samarbeidere vil altså bli større jo mindre K er.¹⁹

¹⁹Dersom $K = 0$ har spillerne full informasjon om hvilken type de møter. Samarbeiderne får $U(S) = \beta$ mens opportunistene får $U(O) = 0$. Siden $U(S) > U(O)$ for alle $p \in (0, 1)$ så vil den evolusjonære prosessen beskrevet ovenfor lede til en situasjon der alle til slutt samarbeider. Dette kan illustrere situasjonen i små tette samfunn der alle kjenner alle.

7 Sekvensielle handlinger

I situasjonen ovenfor er det antatt at partene må velge handling simultant. I mange sammenhenger er dette en antakelse som ikke passer så godt. Det er for eksempel ofte slik at en har mulighet til å observere andres handlinger direkte før en selv velger handling. En slipper da å bruke ressurser på å undersøke hvilken type de andre er siden handlingene deres gir informasjon om dette.²⁰ Den motsatte situasjonen er selvfølgelig også mulig, nemlig at en selv må velge først og deretter de (eller den) andre. La oss se nærmere på hvordan resultatene i avsnitt 6 ovenfor endres dersom vi antar at partene er involvert i et spill med trekkrekkefølge. Situasjonen kan illustreres ved hjelp av figur 3.



Først velger spiller 1 (som kan være en samarbeider eller en opportunist) om han skal samarbeide eller ikke. 'Ikke samarbeid' blir besvart med 'ikke samarbeid' som gir null i utbytte. Dersom spiller 1 velger samarbeid betyr det at han legger ressurser i relasjonen som spiller 2 kan utnytte. Etter at spiller 1 har valgt kan spiller 2 opptre på to måter. Han kan samarbeide og handle i tråd med partenes felles interesser. Da får begge et utbytte på β . Alternativt kan spiller 2 opptre opportunistisk og utnytte det faktum at spiller 1 har gjort en irreversibel investering i relasjonen. Ved å velge 'ikke samarbeid' får spiller 2 et materielt utbytte på α , mens spiller 1 får et tap

²⁰Güth og Kliemt (1998; 2000) analyserer en slik situasjon nærmere. Framstillinger nedenfor bygger på disse arbeidene.

på $-\alpha$.²¹ En opportunist med $\tau = 0$ vil dermed alltid velge ‘ikke samarbeid’ dersom han møter en som allerede har valgt ‘samarbeid’. En samarbeider med $\tau > \alpha - \beta$ vil derimot alltid velge ‘samarbeid’ dersom han møter en som allerede har valgt ‘samarbeid’.

Det som videre skiller situasjonen med sekvensielle trekk fra situasjonen med simultane trekk i avsnitt 6 er at det kan oppstå situasjoner hvor også opportunistene velger ‘samarbeid’ (med simultane trekk er ‘ikke samarbeid’ en dominerende strategi for opportunistene). Dersom en opportunist skal handle først, og han vet at han har å gjøre med en samarbeider, så vil han velge ‘samarbeid’ (siden dette gir β som er et større utbytte enn hva han vil få ved å velge ‘ikke samarbeid’). Vi skal imidlertid holde fast på antakelsen ovenfor om at ingen med sikkerhet vet hvem de møter, med mindre de undersøker (noe som koster K). Det eneste spillerne i utgangspunktet kan observere er andelen samarbeidere (p) i samfunnet.

I situasjonen med simultane trekk analysert ovenfor så vil det aldri lønne seg for opportunistene å bruke ressurser på og undersøke. Hvis $p < p^*$ fra (12) og samarbeiderne undersøker, så vil opportunistene aldri ha mulighet til å møte en samarbeider uansett. Det har dermed ingen hensikt å bruke tid og krefter på å undersøke hvem en møter for opportunistene. Hvis $p > p^*$ og samarbeiderne samarbeider uten å undersøke, så er det bedre for opportunistene å velge ‘ikke samarbeid’ framfor å undersøke. Disse resultatene gjelder ikke i situasjonen med sekvensielle trekk. Med sekvensielle trekk kan det lønne seg for opportunistene (på samme måte som for inspektørene), å undersøke nærmere hvem de spiller mot dersom de blir tillagt rollen som den som skal handle først. Det er dermed bare i det andre trinnet av spillet at samarbeiderne og opportunistene har avvikende adferd. Dersom en samarbeider møter en som samarbeider i første trekk, så vil han svare med ‘samar-

²¹Et eksempel på en situasjon som kan fanges opp av spillet ovenfor, og som er i tråd med tolkningen av spillet i avsnitt 2, er å tenke på spiller 1 enten som en kjøper eller selger av en vare. Dersom spiller 1 er kjøper, så kan følgende utfall inntreffe. For å få tilgang til varen må kjøperen først sende over et avtalt beløp til selgeren. Dersom han ikke sender over beløpet (velger ‘ikke samarbeid’) så vil ikke selgeren gi han varen, og begge får null i utbytte. Dersom kjøperen sender over beløpet (velger ‘samarbeid’) så kan enten selgeren gi han varen (velge ‘samarbeid’), som gir begge et utbytte på β , eller han kan velge å ikke gi han varen (velger ‘ikke samarbeid’). I det siste tilfellet får kjøperen $-\alpha$ mens selgeren får α . Dersom spiller 1 er selger, så kan følgende utfall inntreffe. For at selgeren skal få betaling så må først varen framstilles. Dersom den ikke framstilles (velger ‘ikke samarbeid’) så får han heller ingen betaling, og begge får null i utbytte. Dersom selgeren framstiller varen (velger ‘samarbeid’) så kan enten kjøperen gi han den avtalte prisen (velge ‘samarbeid’), som gir begge et utbytte på β , eller han kan velge å ikke betale for varen (velger ‘ikke samarbeid’). I det siste tilfellet får selgeren $-\alpha$ mens kjøperen får α .

beid'. Dersom han møter en som ikke samarbeider, så vil han svare med 'ikke samarbeid'. Opportunistene på sin side vil svare med 'ikke samarbeid' uansett hva partneren velger i første trekk.

Hva er det som bestemmer hvem av de to typene som skal handle først? Anta at dette bestemmes tilfeldig, det vil si at spillerne havner opp i situasjonen som den som skal handle først med sannsynlighet $1/2$. Ellers er dynamikken i spillet identisk med slik den ble beskrevet i avsnitt 6 ovenfor.

La oss starte med å studere beslutningen til de to typene dersom de ender opp i rollen som den som skal handle først. Ved å velge samarbeid i første trekk vil begge typene få β dersom de møter en samarbeider, noe som vil skje med sannsynlighet p , mens de vil få $-\alpha$ dersom de møter en opportunist, noe som skjer med sannsynlighet $(1-p)$. Dersom de velger 'ikke samarbeid' vil de få 0 uansett. Siden begge typene har informasjon om størrelsen på p , så vil de velge 'samarbeid' dersom $p\beta + (1-p)(-\alpha) > 0$, det vil si dersom

$$p > \frac{\alpha}{\alpha + \beta} \equiv p^\dagger \quad (13)$$

hvor $p^\dagger \in (0, 1)$ siden $\alpha + \beta > \alpha$. Begge typene vil altså velge 'samarbeid' i rollen som den som skal handle først dersom andelen som samarbeider i samfunnet er større enn p^\dagger . Men alle har også mulighet til å undersøke hvilken type de spiller mot før de velger handling, noe som koster K . Anta videre at dersom spillerne velger å undersøke, så må dette gjøres før de vet om de er den som skal handle først eller sist. Spørsmålet er derfor under hvilke betingelser de to typene vil velge å undersøke i første trinn av spillet.

La oss først se på tilfellet hvor $p > p^\dagger$. Når $p > p^\dagger$ vil begge typene, uten informasjon om hvem de spiller mot, velge 'samarbeid' dersom de handler først. Dette gir et forventet utbytte på $\frac{1}{2}[p\beta + (1-p)(-\alpha)] + \frac{1}{2}\beta$. Før de vet om de skal handle først eller sist har de imidlertid mulighet til å undersøke nærmere hvem de spiller mot. Dette gir i så fall et forventet utbytte på $\frac{1}{2}[p\beta + (1-p)0] + \frac{1}{2}\beta - K$. De to typene vil dermed velge å undersøke i rollen som den som handler først dersom $(1+p)\beta - 2K > (1+p)\beta - (1-p)\alpha$, eller

$$p < 1 - \frac{2K}{\alpha} \equiv p^* \quad (14)$$

(14) sier at for en gitt K , så vil de to typene undersøke nærmere hvem de spiller mot dersom $p^\dagger < p < p^*$. Dersom $p > p^*$, så er andelen som samarbeider i samfunnet så stor at det ikke lønner seg å ta kostnaden ved å undersøke. Uavhengig av hvilken type en er, så lønner det seg heller å samarbeide ubetinget dersom en havner i rollen som den som skal handle først.

La oss så se på tilfellet hvor $p < p^\ddagger$. Når $p < p^\ddagger$ vil begge typene, uten informasjon om hvem de spiller mot, velge ‘ikke samarbeid’ dersom de handler først. Før de vet om de skal handle først eller sist har de mulighet til å undersøke nærmere hvem de spiller mot. Ved å undersøke vil de få $\frac{1}{2} [p\beta + (1-p)0] + \frac{1}{2}0 - K$. Ved ikke å undersøke vil de få 0 siden ingen vil samarbeide uansett om de velger først eller sist. De to typene vil dermed velge å undersøke dersom $p\beta - 2K > 0$, eller

$$p > \frac{2K}{\beta} \equiv p^{**} \quad (15)$$

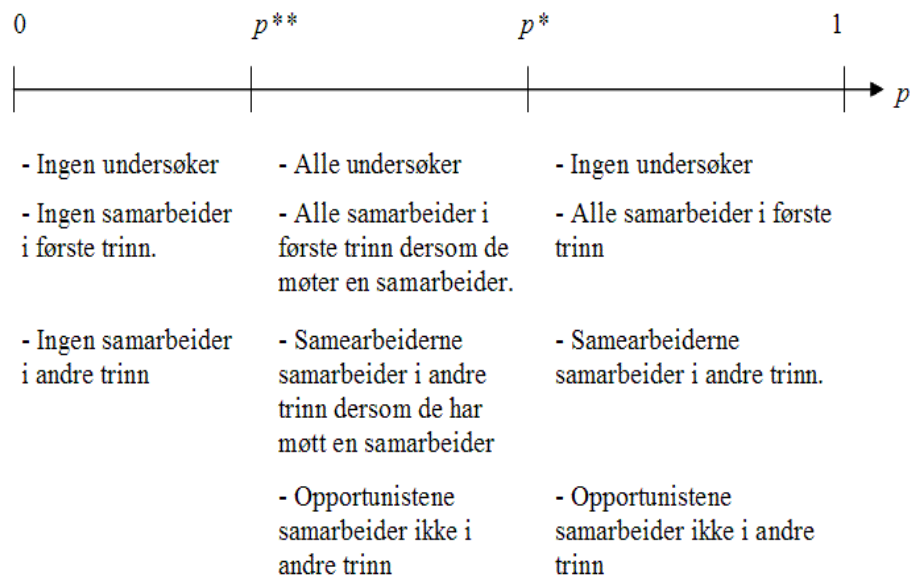
(15) sier at for en gitt K , så vil de to typene undersøke nærmere hvem de spiller mot dersom $p^{**} < p < p^\ddagger$. Dersom $p < p^{**}$, så er andelen som samarbeider i samfunnet så liten at det ikke lønner seg å ta kostnaden ved å undersøke. Uavhengig av hvilken type en er, så lønner det seg heller å velge ‘ikke samarbeid’ dersom en havner opp i rollen som den som skal handle først.

Fra (13) og (14) ser vi videre at p^* avtar og p^{**} øker når kostnaden ved å undersøke øker. Det vil altså si at intervallet for p hvor spillerne tar kostnadene ved å undersøke i første trinn reduseres i takt med at K øker. Det betyr altså at ingen vil ta kostnaden ved å undersøke dersom K blir for stor. Mer presist så må ikke K være så stor at intervallet for p definert i (13) and (14) er tomt, det vil si at

$$K < \frac{\alpha\beta}{2(\alpha + \beta)} \equiv K_{\max} \quad (16)$$

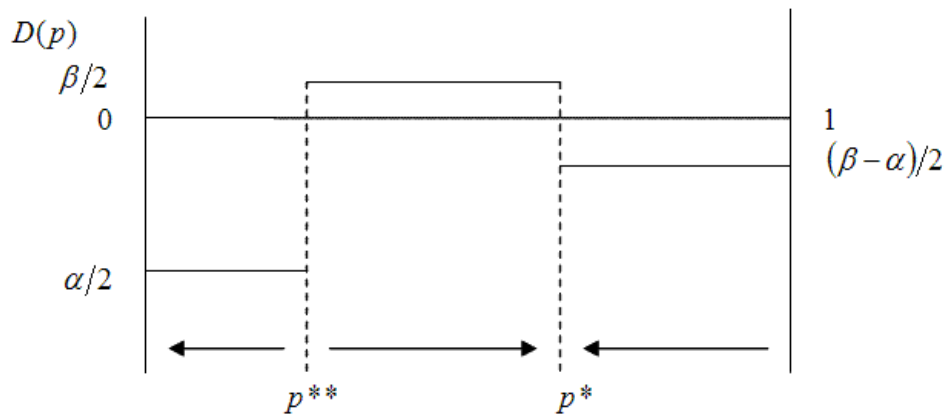
Det er altså bare når kostnaden ved å inspisere er mindre enn K_{\max} at de to typene vil velge å undersøke i rollen som den som skal handle først.²² De to typenes beslutning om å undersøke eller ikke, samt valg av handling i de to trinnene av spillet, for $K < K_{\max}$, er illustrert ved hjelp av figur 4.

²²Dersom $K = K_{\max}$ så vil de to typene bare undersøke dersom $p = \alpha/(\alpha + \beta) \equiv p^\ddagger$. For verdier av p som ligger over eller under p^\ddagger så vil de ikke undersøke.



Figur 4 viser at det bare er i rollen som den som skal handle sist at samarbeiderne og opportunistene har avvikende adferd. Opportunistene samarbeider aldri i andre trinn av spillet, mens samarbeiderne samarbeider dersom de møter en som har samarbeidet i første trinn. Når $p < p^{**}$ vil begge typene velge 'ikke samarbeid' uten å undersøke i rollen som den som skal handle først. Når $p > p^*$ vil begge typene velge 'samarbeid' uten å undersøke, mens de begge vil undersøke og samarbeide dersom de møter en som samarbeider for $p^{**} < p < p^*$.

Hvordan andelen av de to typene utvikler seg over tid når $K < K_{\max}$ kan da framstilles ved hjelp av figur 5.



La oss først se på situasjonen der alle i utgangspunktet samarbeider ($p = 1$), og anta at det av en eller annen grunn dukker opp en liten gruppe opportuniste. Disse vil utnytte de som samarbeider ved å velge ‘ikke samarbeid’ (som gir α) dersom de blir tillagt rollen som den som handler sist (noe som inntreffer med sannsynlighet $1/2$). I rollen som den som skal handle først vil opportunistene (i likhet med samarbeiderne) velge ‘samarbeid’. Opportunistene vil da få et forventet utbytte på $U(O) = (\beta + \alpha)/2$, mens samarbeiderne får $U(S) = \beta$. Benytter vi dette i (3) finner vi at $D(p) = (\beta - \alpha)/2 < 0$ og $\dot{p} < 1$. Opportunistene har et høyere forventet utbytte enn samarbeiderne. Via den evolusjonære prosessen beskrevet i avsnitt 3 vil andelen samarbeidere avta inntil $p = p^*$. Når $p = p^*$ vil samarbeiderne begynne å undersøke hvem de møter. Det vil dermed ikke være mulig for opportunistene å utnytte samarbeidernes samarbeid, siden samarbeiderne i første trekk vil velge ‘ikke samarbeid’ dersom de møter en opportunist. Forventet utbytte til samarbeiderne blir $U(S) = \frac{1}{2}(p\beta + (1-p)0) + \frac{1}{2}\beta$, mens opportunistene får $U(O) = \frac{1}{2}(p\beta + (1-p)0) + \frac{1}{2}0$. Setter vi inn i (3) får vi $D(p) = \beta/2 > 0$ og $\dot{p} > 0$. For $p < p^*$ vil altså samarbeiderne ha et høyere forventet utbytte siden de samarbeider både med opportuniste og andre samarbeidere dersom de tillegges rollen som den som skal handle sist. $p = p^*$ er dermed en stabil indre likevekt.

La oss så betrakte situasjonen der ingen i utgangspunktet samarbeider ($p = 0$), og at det dukker opp en liten gruppe som samarbeider. Disse vil da bli utbyttet av alle de andre medlemmene av samfunnet som ikke samarbeider og får et lavere forventet utbytte. Den lille gruppen samarbeidere vil gradvis endre sin adferd i retning av ‘ikke samarbeid’. Et samfunn bestående utelukkende av ikke samarbeidende opportuniste er med andre

ord en stabil likevekt. Dersom gruppen samarbeidere er så stor at $p > p^{**}$ er imidlertid samarbeid en levedyktig strategi. Når $p > p^{**}$ vil samarbeiderne undersøke hvem de møter. Det vil dermed ikke være mulig for opportunistene å utbytte samarbeidernes samarbeid. Vi er tilbake til situasjonen hvor $D(p) = \beta/2 > 0$ og $\dot{p} > 0$. Straks andelen som samarbeider er større enn $p > p^{**}$ vil altså andelen som samarbeider fortsette å øke inntil $p = p^*$. p^{**} er derfor en ustabil indre likevekt. En liten endring bort fra p^{**} vil føre til en ytterligere bevegelse bort fra p^{**} .

Vi kan dermed konkludere med at det i alt eksisterer fire evolusjonære likevekter i situasjonen ovenfor, nemlig 0, p^{**} , p^* og 1, hvorav den første og den tredje er stabile. Den indre likevekten $p = p^{**}$ definerer et kritisk “vippepunkt” som angir andelen samarbeidere som minst må til for å starte en utvikling i retning av den stabile likevekten p^* . Det betyr altså at samfunnet aldri vil nå en tilstand der samarbeiderne “tar over” hele samfunnet. Samarbeiderne vil heller leve side om side med en andel $(1 - p^*)$ opportuniste. (14) viser imidlertid at den stabile indre likevekten p^* kan ligge nær 1 dersom kostnaden ved å undersøke (K) er liten. Når K er liten ser vi også at det kritiske vippepunktet p^{**} flyttes mot null. Det kreves dermed en mindre andel som samarbeider initialt for å starte en utvikling i retning av den stabile likevekten p^* .²³

Igjen har vi en situasjon der den sosiale interaksjonen kan lede til to svært forskjellige stabile likevekter. Hvor samfunnet ender opp avhenger av hvor det starter, det vil si at den evolusjonære prosessen er *historie-betinget*. Dersom den initiale andelen som samarbeider er liten (mindre enn p^{**}) så vil samfunnet bevege seg i en retning der ingen samarbeider. Dersom den initiale andelen inspektører er stor (større enn p^{**}) så vil samfunnet bevege seg i en retning der flertallet samarbeider. Begge situasjoner er mulig!

8 Avslutning

Muligheten for å skaffe seg informasjon om hvem en møter i en sosial dilemma situasjon kan etablere samarbeid som en stabil løsning dersom en andel av aktørene er utstyrt med en internalisert norm om samarbeid. Dette gjelder selv om det er kostbart å skaffe seg informasjon om hvilken type en samhandler med. Jo større denne kostnaden er, jo vanskeligere er det imidlertid å

²³Vi kan også illustrere dette ved hjelp av et talleksempel. Anta som før at $\alpha = 4$ og $\beta = 2$. Fra (16) får vi da at $K_{\max} = 0.7$. La $K = 0.5$. Fra (14) og (15) får vi da at $p^{**} = 0.5$ og $p^* = 0.75$. Dersom K reduseres til 0.1 reduseres p^{**} til 0.1 mens p^* øker til 0.95.

få etablert samarbeid som en stabil situasjon.

I denne artikkelen er dette resultatet utledet ut i fra ulike antakelser om den underliggende sosiale interaksjonen. Et fellestrekk ved flere av situasjonene som er studert er at den evolusjonære prosessen ender opp i en stabil situasjon der majoriteten samarbeider, men der disse samarbeiderne lever side om side med en andel ikke-samarbeidende opportunister. Dette er et trekk vi finner igjen i de fleste samfunn. Flertallet av borgerne overholder etablerte lover og regler, mens et lite mindretall handler på tvers av dem. Et annet generelt trekk som ser ut til å gjelde er at små stabile samfunn gjennomgående er mindre plaget av opportunistisk adferd sammenlignet med store uoversiktlige samfunn. Sett i lys av modellopplegget ovenfor kan dette forklares ved at kostnaden ved å beskytte seg mot opportunistisk adferd er liten ($K \rightarrow 0$) i små samfunn, slik at gevinsten ved å samarbeide blir stor. I store samfunn, med fravær av tette bånd mellom innbyggerne, øker kostnaden med å skaffe seg informasjon om andre. Dette øker både muligheten til, og utbyttet av, å handle opportunistisk. Graden av asosial adferd øker. Disse resultatene trekker i retning av at det kan være gevinster ved å organisere økonomisk aktivitet i små oversiktlige grupper eller markeder siden dette demmer opp for opportunistisk adferd. Små markeder og liten grad av økonomisk integrasjon reduserer imidlertid mulighetene for spesialiseringsgevinster. Det eksisterer derfor både kostnader og gevinster ved økonomisk integrasjon. Modellopplegget ovenfor er imidlertid for enkelt til å si noe mer presist om størrelsen på gevinster og kostnader av å organisere økonomisk aktivitet innenfor store eller små markeder. Et forsøk på å inkorporere dem vil imidlertid representere en interessant utvidelse av analysen som presenteres i denne artikkelen.

References

- [1] Axelrod, R. and W.D. Hamilton. 1981. The Evolution of Cooperation, *Science*, 211, 1390-1396.
- [2] Basu, K. 1995. Civil Institutions and Evolution: Concepts, Critiques and Models, *Journal of Development Economics*, 43, 19-33.
- [3] Bendor, J. and P. Swistak. 2001. The Evolution of Norms, *American Journal of Sociology*, 106(6), 1493-1545.

- [4] Bowles, S. 2001. Individual Interactions, Group Conflicts, and the Evolution of Preferences. I S. Durlauf og P. Young (red.) *Social Dynamics*. Cambridge: The MIT Press.
- [5] Bowles, S. 2004. *Microeconomics: Behavior, Institutions, and Evolution*. New York: Princeton University Press.
- [6] Bowles, S. og H. Gintis. 2001. *The Economics of Shame and Punishment*. Working paper, Santa Fe Institute.
- [7] Bowles, S. og H. Gintis. 2004. Persistent Parochialism: Trust and Exclusion in Ethnic Networks, *Journal of Economic Behavior and Organization*, 55(1), 1-23
- [8] Boyd, R. og P. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: Chicago University Press.
- [9] Borgstede, von C. 2002. *The Impact of Norms in Social Dilemmas*. Doktorgradsavhandling. Universitetet i Gøteborg.
- [10] Camerer, C. 2003. *Behavioral Game Theory: Experiments on Strategic Interaction*. Princeton NJ: Princeton University Press.
- [11] Coleman, J. 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- [12] Elster, J. 1983. *Explaining Technical Change*. Cambridge: Cambridge University Press.
- [13] Elster, J. 1989. Social Norms and Economic Theory, *Journal of Economic Perspective*, 3(4), 99-117.
- [14] Fehr, E. og S. Gächter. 2002. Altruistic Punishment in Humans. *Nature*, 415(10), 137-140.
- [15] Fehr, E., U. Fischbacher og S. Gächter. 2002. Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms, *Human Nature*, 13, 1-25.
- [16] Frank, R.H. 1988. *Passion Within Reason*. New York: Norton.
- [17] Frank, R.H. 1989. Honesty as an Evolutionary Stable Strategy. *Behavioral and Brain Sciences*, 12(4), 705-706.

- [18] Friedman, J. 1971. A Noncooperative Equilibrium in Supergames, *Review of Economic Studies*, 38, 1-12.
- [19] Fudenberg, D. og E. Maskin. 1986. The Folk Theorem in Repeated Games with Discounting or with Incomplete Information, *Econometrica*, 54, 533-556.
- [20] Gintis, H. 2003. Hitchhikers Guide to Altruism: Genes, Culture and the Internalization of Norms, *Journal of Theoretical Biology*, 206, 169-179.
- [21] Gintis, H, S. Bowles, R. Boyd og E. Fehr. 2003. Explaining Altruistic Behavior in Humans, *Evolution and Human Behavior*, 24, 153-173.
- [22] Güth, W. og H. Kliemt. 1998. The Indirect Evolutionary Approach: Bridging the Gap Between Rationality and Adaption, *Rationality and Society*, 10(3), 377-399.
- [23] Güth, W. og H. Kliemt. 2000. Evolutionary Stable Co-operative Commitments, *Theory and Decisions*, 49, 197-221.
- [24] Guttman, J.M. 2000. On the Evolutionary Stability of Preference for Reciprocity, *European Journal of Political Economy*, 16, 31-50.
- [25] Hechter, M. and K.D. Opp. 2001. What Have We Learned About the Emergence of Norms? I M. Hechter og K.D. Opp (red.) *Social Norms*. New York: Russel Sage Foundation.
- [26] Henrich, J, R. Boyd, S. Bowles, C. Camerer, E. Fehr og H. Gintis. 2004. *Foundation of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.
- [27] Horne, C. 2001. Sociological Perspectives on the Emergence of Norms. I M. Hechter og K.D. Opp (red.) 2001. *Social Norms*. New York: Russel Sage Foundation.
- [28] Marx, K. 1963[1852]. *The Eighteenth Brumaire of Louis Bonaparte*. New York: International Publishers.
- [29] North, D.C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge MA: Cambridge University Press.
- [30] Opp, K.D. 2001. Social Networks and the Emergence of Protest Norms. I M. Hechter og K.D. Opp (red.) 2001. *Social Norms*. New York: Russel Sage Foundation.

- [31] Platteau, J-P. 2000. *Institutions, Social Norms, and Economic Development*. Amsterdam: Harwood Academic Publishers.
- [32] Putnam, R.D. 1993. *Making Democracy Work: Civic Tradition in Modern Italy*. Princeton NJ: Princeton University Press.
- [33] Richerson, P., R. Boyd og J. Henrich. 2003. Cultural Evolution and Human Cooperation. I P. Hammerstein (red.) *Genetic and Cultural Evolution of Cooperation*. Cambridge: The MIT Press.
- [34] Rothstein, B. 2000. Trust, Social Dilemmas and Collective Memories, *Journal of Theoretical Politics*, 12(4), 477-501.
- [35] Sethi, R. 1996. Evolutionary Stability and Social Norms, *Journal of Economic Behavior and Organization*, 29, 113-140.
- [36] Skyrms, B. 1994. Darwin Meets the Logic of Decision: Correlation in Evolutionary Game Theory, *Philosophy of Science*, 61, 503-528.
- [37] Sober, E og D.S. Wilson. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge MA: Harvard University Press.
- [38] Stark, O. 1995. *Altruism and Beyond*. Cambridge: Cambridge University Press.
- [39] Stark, O. 2004. Cooperation and Wealth, *Journal of Economic Behavior and Organization*, 53(1), 109-115.
- [40] Trivers, R.L. 1971. The Evolution of Reciprocal Altruism, *Quarterly Review of Biology*, 46, 35-57.