

Errare humanum est

Det er menneskeligt at fejle. Men dette aspekt af den menneskelige adfærd har traditionelt spillet en helt underordnet rolle i økonomisk teori. I den vanlige udgave er økonomisk teori fejlfri - økonomer kunne måske nok tage fejl, hvad teoriehistorien er fuld af eksempler på - men de af økonomerne beskrevne agenter gør det ikke. Dette er imidlertid blot en konsekvens af, at økonomisk teori har været formuleret ud fra en antagelse om, at agenterne har fuldkommen information, og i så tilfælde er der jo ingen grund til at tage fejl. Men agenterne har faktisk ikke fuldkommen information, hvad en del nyere økonomisk teori da også erkender, og dermed opstår fejlmulighederne. Det er hensigten med dette indlæg at påvise, at fejl og fejlmekanismer spiller en betydelig rolle på såvel den teoretiske som den praktiske plan.

I den statistiske teori er fejlbeslutninger et centralt problem. Der sondres imellem 2 typer af fejl:

Hypotesen burde være	
Hypotesen blev	Accepteret Forkastet
Accepteret	Type II fejl
Forkastet	Type I fejl

*Økonomisk institut
Københavns Universitet*

I praktisk empirisk arbejde bliver interessen i alt væsentligt koncentreret om type II-fejl, f.eks. i form af en (overdreven) interesse for, om en regressor i en regressionsanalyse er signifikant. Allerede Arrow og før ham Wald har imidlertid gjort opmærksom på, at der er omkostninger forbundet med begge typer af fejl¹⁾. Hvilket valg, der bør træffes, må ske ud fra en eksplicit vurdering af konsekvenserne ved at vælge forkert. I nogle tilfælde er konsekvenserne ved den ene type fejl langt alvorligere end ved den anden type af fejl.

Det er tankevækkende, at inden for jura, især strafferetten, har de to typer af fejl været intenst diskuteret. Lad hypotesen være en anklage for en forbrydelse. Det har da været god tone at mene, at type II-fejl - den anklagede blev dømt, selv om vedkommende burde være frifundet - er langt alvorligere end type I-fejl: At frikende en uskyldig. Hele det strafferetlige system er opbygget ud fra denne antagelse. Men måske er der - i hvert fald på et enkelt område - en opblødning på vej af dette synspunkt, nemlig ved den faste promille ved spirituskørsel. Det er muligt, at en bestemt promille ikke gør visse personer uegnede til at køre bil, og de vil dermed med en større promille end tilladt blive »uskyldigt« dømt, men samtidigt opnås at forhindre en væsentlig mængde type I-fejl, eftersom det bliver langt lettere i praksis at få domfældt en række skyldige.

Sådanne overvejelser har imidlertid spillet en underordnet rolle i økonomi,

men i afsnit 2 skal det påvises, at denne debat har et betydeligt budskab til teoretisk interesserede. For at den teoretiske diskussion ikke skal fremstå som noget rent abstrakt er der i afsnit 3 givet en eksemplificering af tankegangen i form af at vise, at vurderingen af en række praktiske kontrolforanstaltninger kan tage sig ganske anderledes ud, hvis de anskues i lyset af en fejlteori. Konkret er censorinstitutionen valgt som eksempel, fordi de fleste har en fornemmelse af dens funktion. I afsnit 4 skal det demonstreres, at fejl og fejlmekanismer også bør interessere empirikere og praktikere, men i stedet for en mere abstrakt diskussion skal problemet behandles ved at gennemgå en ny undersøgelse af AKF om revalidering. Det skal påvises, at konklusionerne fra denne undersøgelse helt afhænger af, hvilke antagelser man gør om fejlmekanismerne ved beslutningen om revalidering.

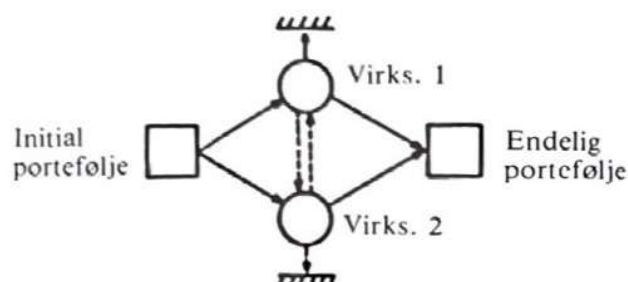
Fejl i økonomiske systemer

I en banebrydende artikel - The Architecture of Economic Systems: Hierarchies and Polyarchies - i septembernummeret af American Economic Review 1985 har Sah og Stiglitz eksplicit analyseret, hvorledes forskellige måder, på hvilken den økonomiske virksomhed kan organiseres, medfører forskellige fejltyper.

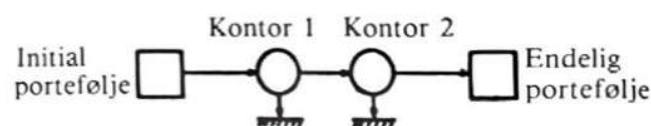
For at nå til mere konkrete resultater betragter de en stiliseret ramme: Lad der være givet en initial portefølje af investeringsprojekter karakteriseret ved deres kapitalværdi x , som kan være positiv eller negativ. De definerer herefter en *evalueringsskærmingsregel* - screening function - $p(x)$ som sandsynligheden for, at et projekt bliver bedømt som godt mod alternativet ikke-godt. Denne sandsynlighed er naturlig nok en stigende funktion af projektets kapitalværdi, gode projekter har en større

sandsynlighed for at blive accepteret, men bedømmelsen af projekterne er imidlertid ikke fuldkommen sikker - i så fald ville problemet være trivielt, vi ville være tilbage i den klassiske verden, hvor alle projekter med $x > 0$ blev i værksat.

Som grundlag for deres analyse betragter de to alternative organisationsstrukturer, et polyarki og et hierarki, se fig. 1:



Figur 1a. Polyarki



Figur 1b. Hierarki

I et *polyarki*, som minder om en markedsøkonomi med mange uafhængige virksomheder, kan et projekt blive accepteret af enten den ene eller den anden virksomhed (men for simpelhedens skyld ikke dem begge). I et *hierarki* er beslutningsprocessen anderledes, thi accept af et projekt kræver accept af både det lavere og det højere kontor. Det lavere kontor vurderer alle projekter, og kun de af det lavere kontor accepterede projekter vurderes af det højere kontor.

Ensartede evalueringsregler. Hvis evalueringen foregår alle steder efter samme principper, d.v.s. hvis $p(x)$ er fælles for alle agenter (kontorer og virksomheder), kan det let vises, at et polyarki udvælger et større antal projekter end et hierarki. Grunden hertil er den simple, at det første kontor hhs. den første virksomhed i henhold til forudsætningen accepterer det samme antal projekter, men virksomhed

2 i et polyarki accepterer i hvert fald nogle af de projekter, som virksomhed 1 har forkastet, mens det højere kontor i et hierarki vil forkaste i hvert fald nogle af de projekter, som det lavere kontor har accepteret. Alle forkastede projekter i et polyarki får en ny chance, mens alle godtagne projekter af det lavere kontor i et hierarki bliver underkastet en ny prøve. Dermed bliver den relative forekomst af type I-fejl større i et hierarki, mens omvendt den relative forekomst af type II-fejl bliver større i et polyarki. Den komparative fordel i et hierarki er at forhindre dårlige projekter, mens et polyarki har sin styrke ved at fremme gode projekter. Hvad der totalt set er at foretrække kommer til at afhænge af den initiale portefølje af investeringsprojekter, som kan varierer fra situation til situation.

Endogene evalueringsregler. Dermed er historien imidlertid ikke slut, men begynder tværtimod at blive spændende, thi *selve den institutionelle ramme, som agenterne handler i, vil påvirke deres adfærd.* D.v.s. at $p(x)$ ikke er den samme i et polyarki og et hierarki. Dette er den centrale pointe. Selve det faktum, at agenten ved, at handlingen foregår i et polyarki eller et hierarki, påvirker beslutningen. En agent i et polyarki ved således, at beslutningen ikke vil blive omgjort andetsteds.

Modelteknisk indfører Sah og Stiglitz, at agenterne observerer, ikke den sande kapitalværdi x , men derimod en beregnet værdi y , som er lige med den sande værdi x tillagt et støjled.

Problemet for agenterne er nu at vælge en *reservationsværdi* R , så alle projekter med observeret $y > R$ accepteres. Reservationsværdien kan således opfattes som et krav til pay-off tid eller alternativt den ønskede interne forrentning. Jo større R -værdi, der vælges, desto større bliver chancen for at undgå dårlige projekter

(d.v.s. type II-fejl), men højere R -værdi øger også risikoen for at gå glip af gode projekter (d.v.s. type I-fejl).

Et centralt resultat hos Sah og Stiglitz er, at $R(P) > R(H)$, reservationsværdien for den enkelte profitmaksimerende virksomhed $R(P)$ i et polyarki er større end reservationsværdien $R(H)$ i et hierarki - til trods for, at målsætningen er den samme for alle agenter, nemlig at maksimere profitten. Begrundelsen er ganske naturlig: I et hierarki ved det lavere kontor, at beslutningen vil blive overvejet endnu engang, mens omvendt det øverste kontor ved, at projekterne har været underkastet en vurdering allerede. I et polyarki ved virksomhederne derimod, at det er alt eller intet. En vigtig praktisk konsekvens er, at private virksomheder garderer sig mod fejltagelser ved at kræve en høj forrentning *ex ante*, mens det offentlige - her opfattet som et hierarki - sætter en lavere reservationsværdi²⁾. *Kravet til rentabilitet i offentlige og private investeringskalkuler behøver med andre ord ikke at være det samme* - en konklusion som dels strider imod gængs tænkning på dette område, dels har vigtige praktiske implikationer.

På det principielle plan er det afgørende, at beslutningsreglerne, adfærden, bliver forskellig på det individuelle niveau (den enkelte virksomhed hhs. kontor). Dette har den helt centrale konsekvens, at man ikke kan foretage sammenligninger mellem 2 organisationsformer uden at tage hensyn til, at agenternes adfærd i de to systemer vil være forskellig - en praktisk illustration af dette resultat gives i næste afsnit. Resultatet betyder således også, at analysen i afsnittet om ensartede evalueringsregler ikke i modsætning til, hvad læseren måske umiddelbart troede, var trivielt, men derimod direkte misvisende, eftersom den opstillede forudsætning ik-

ke i sammenhængen kunne være holdbar.

Det faktum, at agenternes adfærd i forskellige organisationsformer er forskellig, kan meget nemt give anledning til »kontraintuitive« resultater. Sah og Stiglitz viser således i et specialtilfælde, at hvis informationen om projekterne er tilstrækkeligt dårlig, vil et polyarki give et bedre resultat end et hierarki. Hvorfor? Man ville nok umiddelbart have troet, at med meget ringe information ville den dobbelte evaluering i et hierarki netop være værdifuld. Men det kan vises, at evalueringsskriterierne ændrer sig så tilpas meget, at dette ikke er tilfældet.

Med deres bidrag har Sah og Stiglitz på den ene side givet et nyt og godt eksempel på, at studiet af organisationsstrukturer ikke kun er af interesse for sociologer, og på den anden side vist, at det teoretiske studie af komparative økonomiske systemer er andet og mere end blot undersøgelser, af hvilket system der under specielle forudsætninger kan give et pareooptimalt resultat. Men naturligvis er deres analyse baseret på en stiliseret ramme, der i konkrete omstændigheder må tilpasses.

Et tænkt eksempel:

Værdien af censorinstitutionen

Der er tradition i Danmark for, at karaktergivningens kontrolleres af udefra kommende censorer. Dette skal ses som en beskyttelse af de studerende. Men er det nødvendigvis tilfældet? Lad os antage, at der er tale om et etableret studium, hvor der blandt lærere og censorer er enighed om niveauet; det kan f.eks. være et grundfag ved jura- eller økonomistudiet. Det giver derfor mening at anvende begrebet en »korrekt karakter. Bedømmelsen kender naturligvis ikke den korrekte karakter med sikkerhed, men antallet af afgivne karakterforslag vil faktisk ramme den korrekte karakter med en intensitet,

der afhænger af lærerens tidsforbrug. Lad os antage, at læreren med tidsforbruget T i tilfælde af censur har følgende karakterfordeling: $4/6$ af karaktererne er rigtige, men $1/6$ af karaktererne er 1 grad for høje, mens $1/6$ af karaktererne er 1 grad for lave.

For at forebygge disse fejl er der etableret censorordninger. Vi forestiller os, at censor giver karakterer helt uafhængigt af læreren. Men censorer er jo heller ikke fejlfri. Lad os derfor antage, at de har sammen fejlfordeling som lærerne. Der indføres følgende nogenlunde realistiske spilleregler for karaktergivningen: a) Hvis begge foreslår sammen karakter, bliver dette karakteren. b) Hvis der er 1 grad imellem karakterforslagene, vil censor få sit forslag i 50 pct. af tilfældene og læreren sit i de øvrige tilfælde. c) Hvis der er 2 graders forskel, tages den midterste karakter. For 36 studerende vil vi dermed få situationen i tabel 2.

Tabel 2. Lærer.

	Korrekt	For meget	For lidt
Korrekt	16	4	4
Censor for meget	4	1	1
For lidt	4	1	1
I alt	24	6	6

Resultatet af voteringen i tabel 2 bliver f.eks. for kategorien for meget, at 5 får 1 grad for meget, nemlig den ene, hvor begge gav 1 grad for meget, 2 ud af de 4, hvor den ene gav en grad for meget, men hvor den anden gav korrekt, og 2 ud af det tilsvarende tilfælde, ialt 5 tilfælde. Resultatet af censuren kan således sammenfattes til

	Lærer alene	Lærer og censor
Korrekt	24	26
For meget	6	5
For lidt	6	5

Men det er ikke de samme, der får for meget hhs. for lidt efter censuren, eftersom lærerens fejlskøn er blevet korrigeret i $3 + 3 \frac{1}{4} 6$ tilfælde, mens der ved censuren er opstået nye fejl i $2 + 2 \frac{1}{4} 4$ tilfælde. Resultatet af censuren er blevet en moderat forbedring af karaktergivningen - censorinstitutionen synes at virke efter hensigten. Men denne konklusion hænger afgørende på den ofte implicitte forestilling om, at læreren har den samme evaluering med og uden censur, altså netop den forestilling, der blev gjort op med i forrige afsnit. Læreren ved, at der senere skal censureres og kan derfor tillade sig lidt mere forhastede afgørelser i tillid til, at fejl vil blive rettet under censuren.

Lærerens tidsforbrug består i dels selve opgaveretningen, dels i censuren. Hvis det sidste tidselement blev anvendt i stedet på at læse nogle af de vanskeligere opgaver mere omhyggeligt, ville systemet »læreren alene« meget vel kunne tænkes at give et bedre resultat end med censur. Dette ville gælde for de ansvarsbevidste lærere. Men ville mangelen på censur ikke blot friste nogle lærere til at anvende endnu mindre tid på opgaveretningen? Muligvis jo. Men det er dog en overvejelse værd, om dette fænomen ikke kunne imødegås på en anden måde, f.eks. ved en stikprøvecensur kombineret med de eksisterende klageregler. Det nuværende system er ikke fejlfrit, men ganske kostbart.

Et konkret eksempel: Sagsbehandlerfejl og effekten af den danske revalideringsindsats

I empiriske undersøgelser vil forskeren normalt enten formaliseret eller blot i baghovedet have en »model«, ud fra hvilken forskeren baserer sin undersøgelse og drager sine konklusioner. Meget ofte er modellen en ret mekanisk opfattelse af, at agenterne med fuld information maksi-

merer en profit- eller nyttefunktion. Men man kan komme til ganske anderledes resultater, hvis det teoretiske udgangspunkt ikke er baseret på en forestilling om altvindinge og dermed fejlfri agenter. Det følgende er et eksempel herpå.

I en nyligt udkommet og ret omfangsrig rapport fra AKF har Åge Valbak og Jens Wamsler undersøgt virkningerne af revalidering og nået til det resultat, at revalidering er en samfundsmæssigt fornuftig foranstaltning - effekten af revalideringsindsatsen er så stor, at det offentlige får sine penge hjem igen, antageligt mange gange, ved revalideringsforanstaltninger. Problemet er af ret stor betydning, eftersom omkring 39.000 familier i 1984 modtog revalideringsstøtte med et samlet støtbeløb på 880 mill. kr.

Men - desværre - kan der stilles et vist spørgsmålstejn ved denne konklusion³⁾. Problemet er efter min opfattelse, at der ikke er taget tilstrækkeligt hensyn til agenternes, her sagsbehandlernes, manglende information og dermed fejlmuligheder.

I rapporten er der undersøgt 337 personer under revalidering, af hvilke de 62,3 pct. gennemførte det planlagte revalideringsprogram, mens resten ikke gennemførte programmet. Det grundlæggende problem ved denne som ved mange andre effektundersøgelser er at vurdere, hvor mange der ville være blevet selvforsørgende også uden revalidering. Rapporten er da også klart opmærksom på dette⁴⁾. Da der imidlertid ikke umiddelbart foreligger en velegnet kontrolgruppe, har man valgt at sammenholde konsekvenserne af gennemførte revalideringssager med konsekvenserne af afbrudte revalideringssager, som således kommer til at tjene som kontrolgruppe.

Gennemførelsesprocenten var $g = 62,3$, mens selvforsørgelsesprocenterne for hhs.

Tabel 3. 100 personer fordelt på gennemførsel og selvforsørgelse.

	Gennemførte		Afbrudte		I alt Antal
	Antal	Pct.	Antal	Pct.	
Selvforsørgende.....	52,8	84,7	15,6	40,5	68,4
Ikke-selvforsørgende	9,5	15,3	22,1	59,5	31,6
I alt antal	62,3		37,7		100,0
Prc.		100,0		100,0	

de gennemførte og de afbrudte revalideringsprogrammer var $s = 84,4$ og $a = 40,5$. I tabel 3 er disse hovedtal angivet.

Den samlede effekt E af revalideringen bliver antal selvforsørgende ved revalidering fratrasket antal personer i en kontrolgruppe, hvor man altså sætter selvforsørgelsesgraden i kontrolgruppen lige med selvforsørgelsesgraden blandt de afbrudte revalideringssager. Vi får dermed:

$$E = 62,3 (84,7 - 40,5) = 27,5$$

Omkring hver fjerde revalideringssag resulterer således i, at den pågældende bliver reelt selvforsørgende. Rapporten tolker reelt disse tal derhen, at selv om der er et ret betydeligt »spild« undervejs, kan de vellykkede revalideringer let betale omkostningerne ved spildet og mere til.

Men hele denne konklusion afhænger af, at der ikke optræder sagsbehandlerfejl, hvad der er vanskeligt at forestille sig på dette område. Lad os definere en *oplagt sagsbehandlerfejl* som det tilfælde, hvor den pågældende klient aldrig nogensinde burde have fået tilbud om revalidering, da det var totalt udsigtsløst, at den pågældende klient skulle kunne gennemføre et revalideringsprogram, endsige få arbejde bagefter. Den slags fejl burde naturligvis være undgået, men optræder utvivlsomt alligevel. Lad der være x procent af den slags tilfælde. Hvis fejlene ikke var forekommet, ville vi i stedet for tabel 3 have observeret tabel 4:

Tabel 4. Revalideringens effekt uden oplagte sagsbehandlerfejl.

	Gennemførte	Afbrudte
Selvforsørgende	52,8	15,6
Ikke-selvforsørgende ..	9,5	22,1 - x
I alt	62,3	37,7 - x

Gennemførselsprocenten bliver nu $62,3/(100-x)$ og selvforsørgelsesgraden blandt de afbrudte revalideringsforløb $15,6/(37,7-x)$. Sættes f.eks. $x = 15$, fås den samlede effekt

$E = 62,3 (84,7 - 15,6/(37,7-15)) = 12,2$ altså en væsentlig reduktion i forhold til de 27,5 i rapporten. Der er ingen tvivl om, at rapportens konklusioner er følsomme overfor forkomsten af oplagte sagsbehandlerfejl. Faktisk vil med den anvendte beregningsmetode revalideringen komme til at udvise et *negativt* resultat, hvis der er over 19,2 pct. oplagte sagsbehandlerfejl i materialet⁵⁾. Jo flere sagsbehandlerfejl, der er inkluderet i materialet, desto større bliver den beregnede effekt af revalideringen. *Den anvendte beregningsmetode medfører således, at jo dårligere, socialrådgiverne har udført deres arbejde, desto bedre vil på papiret revalidering tage sig ud.*

Der er ingen tvivl om, at revalideringssager er vanskelige. Sandsynligheden for, at klienten gennemfører et revalideringsprogram, er således stærkt klientspecifik.

Motivation og psykiske faktorer spiller en stor rolle: Kun i 5 ud af 31 sager, hvor klientens primære ønske om revalidering ikke blev gennemført (eller hvor klienten ikke havde noget formuleret ønske), blev revalideringen gennemført (p.133). Efter omstændighederne kan det nok betegnes som en oplagt sagsbehandlerfejl at forsøge at revalidere lidet motiverede personer. Det samme synes at gælde personer, der kun bliver tilbudt et institutionsophold uden andre foranstaltninger, måske mere for at slippe af med vedkommende end p.g.a. udsigten til revalidering⁶).

Sagsbehandling på et så vanskeligt område som revalidering må naturligvis begå fejl, og selve det faktum, at gennemførselsprocenten ligger væsentlig under 100, indikerer ud fra et bureaukratisk synspunkt dette. Men heri er der intet kritisabelt. Problemet opstår imidlertid, når man anvender en beregningsmetode, hvor sagsbehandlerfejl kan påvirke resultatet ganske væsentligt. Jeg ved naturligvis ikke, hvor stort antallet af oplagte sagsbehandlerfejl er, men det er næppe nul.

Alternativt til den noget håndfaste formulering her kunne mere raffinerede formuleringer af sagsbehandlerfejl udformes. Men den her anvendte formulering er forhåbentlig nok til at demonstrere, at når der er tale om vanskelige beslutninger p.g.a. mangelfuld information, kan agenternes (sagsbehandlerne) adfærdsmønstre ikke negligeres. Det er det i denne rapport, men forhåbentlig ikke i den næste. Kritik skulle nødtigt føre til, at der ikke foretages forsøg.

Det er også en fejl ikke at begå fejl

Det er min opfattelse, at der med den organisatoriske opbygning af i hvert fald den offentlige, men nok også store dele af den private sektor er solide incitamenter

til, at der ubevidst begås langt færre fejl af type II end af type I. Lad os tage revalidering fra forrige afsnit som eksempel. Man kan måle effekten af revalideringsforanstaltninger, udføre analyser af gennemførselsprocenter og måske sågar rangordne de individuelle sagsbehandlere efter, hvor succesrige de er i deres anbefalinger. Den slags analyser udføres der efterhånden en hel del af. man får tabellagt og diskuteret type II-fejlene. Men hvad man typisk ikke interesserer sig nær så meget for, er antallet af personer, der burde revalideres, men ikke bliver det - type I-fejlene. Den slags indgår ikke som led i nogen styrelses rutinemæssige statistik eller årsberetning, men belyses højst en gang imellem som led i et forskningsprojekt. på universiteterne bruges megen tid på at vurdere, om de tilmeldte bør bestå eksamenerne eller ej, men der bruges meget lidt tid på at diskutere, om der er store uudnyttede mængder af talent, som burde være tilmeldt en bestemt videregående uddannelse, men som ikke af den ene eller den anden grund bliver det. Der er ikke i systemet indbygget noget incitament til at interessere sig for den opgave.

Det er derimod for den enkelte medarbejder eller styrelse ofte af afgørende betydning at undgå type II-fejlene. En patient, som blev opereret, men som ikke burde være blevet det, kan lave skandale, anlægge sag mod læge/hospital o.s.v. En minister, der er ansvarlig for et uheldigt projekt kan komme til at lide den tort at blive konfronteret med en forhørsdommer eller en kommissionsdomstol. Det hører unægteligt til sjældenhederne, at ministre eller embedsmænd får sådanne ubehageligheder ved ikke at have iværksat gode ideer. Også forskere kender til at blive hårdere kritiseret for fejl i publikationer end for at undlade at publicere inspirerende nye tanker.

Men hvis man erkender, og det gør man med varierende sprogbrug jævnt hen ved festlige lejligheder, at der burde ofres noget mindre tid på at undgå type II-fejlene og satse noget mere på nye tanker for at mindske type I-fejlene, får dette også en række konsekvenser både i universitetsverdenen og i den praktiske administration. Lad mig give et par eksempler ud af mange mulige.

- I universiteternes økonomiundervisning bruges der f.eks. en rimelig tid på at undervise i investeringsteori og cost-benefit analyse, hvad der i denne sammenhæng er at betragte som lærebøger til elimination af type II-fejl. Tilrettelæggelse af informationssystemer, så type I-fejl kan undgås, betragtes af mange som nærmest værende et problem af ikke-økonomisk karakter. Men set fra virksomhedernes synspunkt er det ikke oplagt, hvilken type fejl det er vigtigst at undgå.

- Der har for nylig været en debat om vejrtjenesten, hvor man ud fra traditionelle effektivitetshensyn har foreslået de 3 eksisterende vejrtjenester sammenlagt. Men dette er ikke kun et effektivitetsproblem. En sammenlagt vejrtjeneste fungerer ikke som 3 uafhængige vejrtjenester, blot lidt billigere. Tilsvarende tanker kan gøre på mange andre områder, f.eks. den stående debat inden for forsvaret om et enhedsværn kontra 3 uafhængige værn.

Både administratorer og forskere burde gøre sig klart, at med valg af organisationsstruktur vælges der også mellem de to fejltyper. Men et sådant synspunkt er måske nok lidt for radikalt i første omgang. Til en begyndelse ville det nok være et afgørende fremskridt, hvis man ville anerkende fejl som en naturlig ting.

Slutnoter

1. Se Kenneth J. Arrow: »Decision Theory and the Choice of a Level of Significance for the t-Test« i Ingram Olkin m.fl. (red.): *Contributions to probability and statistics - Essays in Honour of Harold Hotelling*. Stanford University Press, Stanford California, 1960.
2. En uddybning heraf kan findes i mit memo »Nogle praktiske problemer i cost-benefit analyser«. *Cykelfdelingens memo nr. 72*. Københavns Universitets økonomiske Institut. 1979. Det er med en vis tilfredsstillelse, at jeg noterer mig, at mine synspunkter om kalkulationsrentefoden således efterfølgende har fået en vis opbakning.
3. Det skal understreges, at denne rapport - Åge Valbak og Jens Wamsler: *Revalideringens erhvervsmæssige effekt*, AKF-forlag 1986 - bliver diskuteret for at præcisere en principiel problemstilling. Der er ikke tale om en anmeldelse eller en vurdering af rapporten iøvrigt.
4. Jfr. således følgende citater: »Der vil dog stadig kunne være tale om en tilnærmelse til en ideel effektmåling, idet der må formodes at være forskelle mellem klienter, der kan gennemføre en revalideringsproces, og klienter, der ikke kan gennemføre en revalideringsproces, som ikke kan måles i baggrundsvariable.« (p.71). Og senere p. 114: »Betingelserne for en sådan præcis måling af revalideringens effekt vil formodentlig aldrig være til stede. Det nærmeste, man i praksis kan komme sådan en situation, er at sammenligne den erhvervsmæssige status for klienter, der gennemfører en revalidering, med den erhvervsmæssige status for klienter, der får tilbudt, men ikke gennemfører en revalidering. En sådan sammenligning vil i det følgende være udgangspunktet for vor vurdering af revalideringens effekt.« Se også p. 117.
5. Det kritiske antal oplagte sagsbehandlerfejl, hvor selvforsørgelsesprocenten for de to grupper bliver lige store, fås ved $15,6/(37,7 - x) \frac{1}{4} 0,847$ eller $x \frac{1}{4} 19,2$.
6. Med rapportens ord (p. 138): »Man kan sætte spørgsmålstegn ved det hensigtsmæssige i, at institutionsophold, når de benyttes, tilsyneladende ofte bruges for at fastslå, at reelle revalideringsmuligheder ikke er til stede.«