

Hvem er redd for Homo Oeconomicus?

Om eksperimenter, hevntvang, og krypdyr¹

Karine Nyborg

Født 1962

Professor ved Økonomisk institutt, Universitetet i Oslo

Penger er ikke alt. Nå har selv økonomene skjønt det.

Se for deg følgende: En mann ringer barneavdelingen der han, på bakgrunn av falske eksamenspapirer, er ansatt som sykehuslege. Han har omgangssyke, forteller han, og må holde seg hjemme. Han går deretter ut, praier en drosje og ber om å bli kjørt til Oslo City, der han forlater bilen, uten å betale, blander seg med mengden, finner en kafé og bestiller en bedre lunsj, som han nyter mens han dikter opp forskningsresultatene han skal presentere i sin neste vitenskapelige artikkel. Så tar han bestikket og lysestaken med seg og går, uten å legge igjen en krone. På veien hjem, idet han passerer naboens hage, stikker han innom og tar med seg skjortene som henger til tørk i sola. En helt vanlig morgen, altså, i livet til *Homo Oeconomicus* – økonomifagets mest brukte modell for menneskelig motivasjon og atferd.

Om alle var slik, ville samfunnet åpenbart sett annerledes ut enn det vi er vant til. Ingen skjorter ville hengt til tørk ute. Taxisjåfører, passasjerer, servitører og kafégjester ville brukt store ressurser på å kontrollere hverandre, eller rett og slett holdt seg unna handel; og ingen myndighet ville finne på å tilby offentlig finansiering av egenmeldt sykefravær.

Hvem er Homo Oeconomicus?

Homo Oeconomicus er den latinske hankjønnsformen for *det økonomiske mennesket*.² Vanligvis brukes uttrykket om den enkleste lærebokfremstillingen der sentrale antakelser er perfekt rasjonalitet, perfekt egoisme, og perfekt selvkontroll. Den første som dukket opp da jeg bildegooglet *Homo Oeconomicus* var Skrue McDuck, en illustrasjon som ikke er så verst. *Homo Oeconomicus* bryr seg kun om sin egen tilgang på private og kollektive goder. Han er rasende intelligent, og kan gjennomføre de mest kompliserte beregninger uten anstrengelse. Han har enestående selvkontroll og evne både til å legge, og følge, langsiktige planer. I tillegg er han kynisk kalkulerende og bruker alle

¹ Artikkelen er dels basert på et foredrag i anledning Samfunnsøkonomenes forenings 100-årsjubileum høsten 2008.

² *Oeconomicus* er det latinske ordet for *økonomisk*. Likevel brukes ofte uttrykket *Homo Economicus*, uvisst av hvilken grunn, i den engelskspråklige faglitteraturen (jeg har selv bidratt til dette, før jeg visste bedre). For øvrig begrenser min kompetanse i latin seg til å lytte til råd fra kollega Aanund Hylland.

disse evnene til konsekvent å fremme sine egne interesser.³ Riktignok er han ikke bare opptatt av det materielle. Han vil gjerne ha et stabilt klima, globalt biomangfold, og fred. Problemet er at han aldri foretar seg noe som koster mer enn det smaker for ham selv; derfor vil han aldri kjøpe klimabilletter når han skal ut og fly, siden han alternativt kunne brukt pengene på noe morsommere, eller nyttigere, og selv neppe ville merket noen reduksjon i global temperatur som følge av eget bidrag. Han er gratispassasjer, og håper andre ordner opp. Kjærlighet, hat, vennskap, sosial status eller moralsk plikt er fremmede begreper for ham. Han er *asosial* og *amoralisk* – ikke *usosial* eller *umoralisk*, han bryr seg bare ikke. Hvis *Homo Oeconomicus* skulle forville seg inn på et psykiaterkontor, ville han antakelig blitt tildelt en relativt alvorlig diagnose: dyssosial personlighetsforstyrrelse fremstår som et nærliggende alternativ.⁴

Hva skal vi med ham?

Det er fristende å gjøre narr av *Homo Oeconomicus*, i all sin ufordragetlighet er han det perfekte mobbeoffer – og ikke bryr han seg om at han blir mobbet heller. *Homo Oeconomicus* er en karikatur. Økonomene vet at han er en karikatur, han var aldri ment som annet enn en karikatur, og som sådan er han usedvanlig enkel å ha med å gjøre. Enkelhet er en dyd i forskning: En argumentasjonsrekke bestående av veldige antall ledd, sidespor og komplikasjoner vil, for de fleste av oss, ikke bidra til økt forståelse – halvveis i resonnementet har vi forlenget glemt hva vi begynte å snakke om. Riktignok har ikke *Homo Oeconomicus*, med sine ubegrensede kognitive evner, dette problemet selv, men i og med at økonomer flest er vesentlig mindre begavet enn ham, må vi forenkle. Og skal vi studere etterspørselen etter spiker, for eksempel, spiller det neppe særlig stor rolle hva vi antar om egoisme eller viljestyrke – da kan det enkleste være det beste.

En karikatur kan påpeke viktige trekk som ellers kunne gått oss hus forbi. Det langvarige samlivet med *Homo Oeconomicus* har utvilsomt gitt økonomer god trening i å forstå mange viktige samfunnsmekanismer: som hvordan velmente velferdsordninger kan misbrukes, eller gi uheldige og utilsiktede atferdsendringer, og hvordan miljøavgifter og omsettelige utslippskvoter effektivt kan få selv den mest skruppelløse gratispassasjer til å bli miljøvennlig.

Til tider kan *Homo Oeconomicus* også oppføre seg ordentlig, ikke minst når han har et rykte å ta vare på. Likevel er det vanskelig å komme forbi at karikaturer også kan bidra til at vi glemmer trekk tegneren har utelatt. *Homo Oeconomicus* har vært atskillig mindre til hjelp når det gjelder å forstå hvorfor mange *ikke* misbraker velferdsordningene. Og hva er det som faktisk *får* enkelte flypassasjerer til å kjøpe klimabilletter? Mer generelt: Hvorfor er det slik at mennesker tidvis, men ikke alltid og overalt, greier å opprettholde samarbeid og tillit, selv uten tvang eller økonomiske sanksjoner? Dette er et grunnleggende spørsmål, der vår forståelse fremdeles er sørgelig ufullstendig. Når det gjelder klimaproblemet, har de fleste analyser med utgangspunkt i *Homo*

³ Spesialservice for økonomer: Altså, $\max U_i = u_i(c_i, G)$, der $c_i = i$'s tilgang til private goder, $G = i$'s tilgang til fellesgoder, u_i er en nyttefunksjon som er kontinuerlig, stigende, strengt kvasikonkav og to ganger kontinuerlig deriverbar i begge variable. (Ble det klarere nå?)

⁴ "[P]ersonlighetsforstyrrelse karakterisert ved likegyldighet for sosiale forpliktelser og hensynsløs mangel på forståelse for andres følelser [...] Det finnes ingen effektiv behandling." (Store Norske Leksikon, http://www.snl.no/dyssosial_personlighetsforstyrrelse).

Oeconomicus konkludert med at det ikke er mulig å få til stabile, internasjonale klimaavtaler med flere enn maksimalt tre parter.⁵ Til tross for disse nedslående resultatene, fortsetter verdens ledere å forhandle; og kanskje finnes det likevel utveier, dersom vi tar hensyn til at klimaavtaler inngås av *Homo Sapiens* og ikke *Homo Oeconomicus*. Men om så er tilfelle, og hvilke betingelser det eventuelt krever, vet vi – etter tiår med miljøøkonomisk forskning – fremdeles svært lite om.

Spare- og forsikringsatferd, fedme og rusavhengighet har ofte vært studert ved hjelp av *Homo Oeconomicus*-modellen. Problemet er at disse analysene forutsetter perfekt selvkontroll, og rimeligheten av dette kan i høyeste grad diskuteres. Finanskrisen har synliggjort at *Homo Oeconomicus*-modellen har sine klare begrensninger også i studier av finansmarkedet: En viktig årsak til krisen var fremveksten av nye finansinstrumenter som var så kompliserte at selv profesjonelle investorer lot seg forlede til grovt å undervurdere risikoen, og det i en slik grad at markedet var ute av stand til å korrigere feilene før omfanget ble formidabelt. Om markedsaktørene hadde hatt *Homo Oeconomicus*' kognitive evner, ville krisen neppe oppstått. For å få en bedre forståelse av markeds mekanismer som potensielt kan føre til full systemkollaps, trenger vi kunnskap om hvordan kognitive begrensninger kan påvirke spekulative markeder. Summa summarum: Karikaturen er nok nyttig, men egner seg ikke til alt.

Atferdsøkonomi og eksperimenter

Mange, også økonomer, har kritisert økonomifaget for å være ensidig bygget opp omkring *Homo Oeconomicus*-modellen. De siste femten årene eller så har imidlertid den faginterne kritikken fått en særdeles konstruktiv form, og et helt nytt forskningsfelt, atferdsøkonomi, har vokst fram. Atferdsøkonomien er kjennetegnet nettopp ved systematiske studier av avvik fra, eller utvidelser av, *Homo Oeconomicus*-modellen. Det kunnskapsmessige fundamentet for disse utvidelsene er ofte hentet fra andre fag, særlig psykologi, men også blant annet antropologi, sosiologi, biologi og nevrologi.

Atferdsøkonomien har hatt en rivende faglig utvikling: Fra å være et marginalet forskningsfelt for spesielt interesserte, er feltet allerede langt på vei akseptert som en integrert del av økonomifagets kjerne. Nye tanker trenger riktignok tid til spredning. Fortsatt er de fleste økonomer med ansvar for politikk og byråkrati utdannet på et tidspunkt da få eller ingen hadde hørt om atferdsøkonomi. På forskningsfronten er det imidlertid liten tvil om at atferdsøkonomien har fått bred innpass, også ved de internasjonalt mest prestisjefylte universitetene og i de beste tidsskriftene.

Det er likevel ikke slik at atferdsøkonomiske analyser bygger på et helt annet teorigrunnlag, eller helt andre metoder, enn økonomisk forskning for øvrig. Som regel tar atferdsøkonomiske modeller utgangspunkt i tradisjonell økonomisk teori, men modifierer og/eller utvider den på enkelte punkter. Når jeg jobber med teoretisk analyse av moralnormer, er det for eksempel naturlig å starte med *Homo Oeconomicus*, og deretter utvide modellen ved å inkludere en mulig presisering av hva "moralisk motivasjon" kan bety. En god forståelse av *Homo Oeconomicus* – personlighetsforstyrret eller ikke – er derfor viktig også for atferdsøkonomer: Ikke fordi vi tror han representerer en gjennomsnittsperson, men fordi han utgjør et nyttig og enkelt utgangspunkt for analysen.

⁵ Se for eksempel Barrett, S. (1994): Self-Enforcing International Environmental Agreements, *Oxford Economic Papers* 46, 878-894.

Atferdsøkonomien representerer derfor ikke en helt ny type økonomisk teori, men snarere en videreutvikling av eksisterende tradisjon – og kanskje er dette nøkkelen til gjennomslagskraften feltet har vist seg å ha. At slik videreutvikling av og til kan ha stor betydning for konklusjonene som trekkes, er en annen sak.

Utviklingen av atferdsøkonomi har gått hånd i hånd med framveksten av *eksperimentell økonomi*, et felt som i løpet av 2000-tallet har hatt en nærmest eksplosiv vekst. Poenget med et økonomisk eksperiment er å undersøke hvordan menneskelig atferd varierer med ulike insentiver og omstendigheter. De kontrollerte forholdene i et laboratorium gjør det mulig å endre én faktor om gangen, mens omgivelser og insentiver for øvrig holdes uendret. Økonomiske eksperimenter skiller seg fra psykologieksperimenter blant annet ved at deltakerne så å si alltid betales i penger, der betalingen til den enkelte avhenger av hvilke valg han/hun, ofte også andre deltakere, tar underveis i eksperimentet.⁶

Vi er mer sjenerøse, og mer hevnjerrige, enn Homo Oeconomicus

Noen typer eksperimenter er gjentatt mange ganger, under vidt forskjellige forhold, og med forbløffende like resultater. La meg nevne noen eksempler.

Diktatorspillet er et usedvanlig enkelt eksperiment. Deltaker A får tildelt et beløp, la oss si 100 kroner, og får beskjed om å fordele disse mellom seg selv og en annen deltaker, B. Ingen av dem får vite hvem den andre er. *Homo Oeconomicus* vil selvsagt, i rollen som A, ta alt selv. Et utall eksperimenter viser imidlertid at ikke alle er som ham: I snitt gir A omtrent 20 prosent til B. Å ta alt selv er vanlig, men det er også svært vanlig å dele likt.⁷

Mens B kun er passiv mottaker i diktatorspillet, har han valget mellom å takke ja eller nei i *ultimatuspillet*. Sier han ja, får begge det A har foreslått, sier han nei, får ingen av dem noe i det hele tatt. Hva predikerer *Homo Oeconomicus*-modellen her? Jo: B vil aldri si nei til en positiv sum penger, derfor vil han/hun alltid foretrekke den minste lille sum framfor ingenting, og ta til takke med 50 øre; og fordi A vet dette, og vil ha mest mulig selv, vil A aldri tilby mer enn 50 øre.

I laboratoriet finner man typisk at A tilbyr omtrent 40 prosent av beløpet de er tildelt, altså mye mer enn *Homo Oeconomicus*-modellen predikerer. 50 prosent er svært vanlig. Vel så interessant – og her ligger kanskje noe av forklaringen på det første – er imidlertid at lave tilbud ofte avslås: Hvis A tilbyr mindre enn 20 prosent, vil B avslå i omtrent halvparten av tilfellene.⁸

Ultimatuspillet er studert over hele verden, i mange ulike kulturer, og med opptil flere månedslønner som innsats. Mens resultatene bare i liten grad varierer med innsatsens størrelse,

⁶ I tillegg er det å *lure* eksperimentdeltakerne fullstendig tabu i økonomiske eksperimenter; i psykologi regnes dette, under visse, strenge forutsetninger, som akseptabelt.

⁷ Camerer, Colin (2003): *Behavioral Game Theory. Experiments in Strategic Interaction*. Princeton University Press/Russell Sage Foundation.

⁸ Ibid.

varierer de til en viss grad med kultur.⁹ I samfunn der det å gi store gaver er et tegn på sosial overlegenhet, er det for eksempel ikke uvanlig at A gir mer enn 50 prosent, og at slike forslag avslås!¹⁰ En kultur der folk i det store og hele opptrer som *Homo Oeconomicus*, har man likevel fortsatt til gode å identifisere. Den gruppen som – basert på ulike typer eksperimenter – kan se ut til å opptre mest som *Homo Oeconomicus*, er barn i 3 – 4 års alder¹¹.

Etterhvert som den eksperimentelle litteraturen har vokst, synes visse trekk ved menneskelig atferd å avtegne seg nokså gjennomgående, på tvers av type eksperiment. Ikke minst deler vi langt mer med hverandre, og bidrar betydelig mer til fellesgoder, enn *Homo Oeconomicus*. Samtidig straffer vi hverandre langt mer enn *Homo Oeconomicus* ville gjort. Hvis vi føler oss dårlig behandlet, tar vi igjen så det synger – og dette gjør vi selv om det koster noe og vi ikke har en krone å tjene på avstraffelsen.

Dette kan illustreres ved *fellesgodespillet*: Deltakerne deles i grupper på f.eks. fire, men får ikke vite hvem de andre medlemmene er. Hver av dem skal så – samtidig, og anonymt – dele et beløp, la oss si 100 kroner, mellom seg selv og gruppa. Alt en person tar selv, får hun med hjem; alt hun gir til gruppa, fordobles og deles likt på alle gruppemedlemmene. I et slikt spill lønner det seg for den enkelte å ta hele hundrelappen selv, men for gruppa er det best om alle bidrar med alt. I praksis ser vi at folk i snitt bidrar 40 – 60 prosent; en del bidrar ingenting, mange bidrar noe, og noen gir alt. Dersom et slikt spill gjentas flere ganger, ser vi imidlertid som regel at bidragene faller dramatisk over tid, selv når folk stadig plasseres i nye grupper. Hvis høye bidragsyttere havner i grupper med hverandre, klarer de derimot å opprettholde samarbeidet.¹²

Mye tyder på at både sjenerøsiteten og hevnjerrigheten kan knyttes til det som kalles *resiprositet*,¹³ eller gjensidighetspreferanser: et ønske om å gjengjelde godt med godt og ondt med ondt. Merk at dette dreier seg om en preferanse (hvis jeg tror du utnytter meg, får jeg lyst til å ta igjen), ikke en strategi (jeg tror det er lurt å ta igjen). Folk med slike preferanser vil tendere til å være betingede bidragsyttere: De bidrar mer til fellesskapet jo mer andre bidrar, men trekker seg tilbake hvis de blir skuffet over andres innsats. Dette synes faktisk å være svært vanlig, og muligens vanligere enn *Homo Oeconomicus*-preferanser: For eksempel fant Urs Fischbacher og Simon Gächter i et

⁹ Ibid.

¹⁰ Heinrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, og R. McElreath (2001): In Search of Homo Economicus: Behavioral Experiments in 15 small-scale societies, *American Economic Review* **91** (2), 73-78.

¹¹ Se Fehr, E., H. Bernard og B. Rockenbach (2008): Egalitarianism in young children, *Nature* **454**, 1079-1084, og Ingvild Almås, Erik A. Sørensen og Bertil Tungodden (2009): Development of fairness preferences in children, NHH: foreløpig upublisert notat.

¹² Se f.eks. Ledyard, J.O. (1995): "Public Goods: A Survey of Experimental Research", i J.H. Kagel and A.E. Roth (red.): *The Handbook of Experimental Economics*, Princeton: Princeton University Press; Fehr, E., og S. Gächter (2002): Altruistic Punishment in Humans, *Nature* **415**, 137-140; eller Gächter, S. og C. Thöni (2005): Social Learning and Voluntary Cooperation among Like-Minded People, *Journal of the European Economic Association* **3**(2-3), 303-314.

¹³ Rabin, M. (1993): Incorporating Fairness into Game Theory and Economics, *American Economic Review*, **83** (5), 1281-1302.

fellesgodeeksperiment at 55 prosent av deltakerne kunne klassifiseres som betingede bidragsytere, mens bare 23 prosent oppførte seg som *Homo Oeconomicus*.¹⁴

Gjengjeldere, klima, og velferdsstat

I et perfekt frikonkurransemarked – det vil si en hypotetisk verden der alle goder kan stykkes opp og selges, der alt av betydning reflekteres i markedspriser og der verken maktutøvelse eller lureri er mulig – finnes få muligheter til å belønne og straffe andre. I et slikt marked ville man ikke nødvendigvis merke noen særlig forskjell på gjengjeldere og *Homo Oeconomicus*. I en verden der enkeltmennesker faktisk påvirker hverandre, på godt og vondt, stiller saken seg annerledes.

Mennesker med gjensidighetspreferanser – la oss kalle dem gjengjeldere – har et større potensiale for samarbeid enn *Homo Oeconomicus*. *Homo Oeconomicus* er bare villig til å ofre noe på samarbeidets alter dersom det lønner seg for ham selv og han ikke har mulighet til å være gratispassasjer. En gjengjelder, derimot, kan godt foretrekke å gjøre et offer for samarbeidets skyld, såfremt han føler at andre også legger godviljen til. Kanskje er det gjengjelderne vi trenger for å oppnå en stabil internasjonal klimaavtale: Hvis partene i avtalen er gjengjeldere, kan de være villige til å bidra selv om de strengt tatt kunne kommet enda bedre ut ved å være gratispassasjerer på andre lands innsats.¹⁵

Det er verdt å merke seg at også *Homo Oeconomicus'* atferd vil påvirkes av at gjengjelderne finnes i samfunnet: Hvis du snyter tannlegen din, for eksempel, og han viser seg å være en gjengjelder, kan du fort komme til å angre. Selv en kynisk egoist vil tenke seg om to ganger før han fornærmer en annen, dersom den andre muligens er av den hevnjerrige sorten. Gjengjeldernes potensielle hevntørst – og også deres potensielle belønningsiver, hvis man behandler dem bra – disiplinere også de skruppelløse.

For di gjengjeldere ikke bare er (betinget) sjenerøse, men også hevnjerrige, har de imidlertid også et større potensiale enn *Homo Oeconomicus* for virkelig å lage kvalm. *Homo Oeconomicus* ville aldri funnet på å ødelegge bedriftens dataarkiv siste dag på jobb, true med å skyte sjefen, eller endog gjøre det, for å hevne at han er blitt sagt opp. En gjengjelder, derimot, kunne være troende til slikt.¹⁶ En befolkning med en stor andel gjengjeldere vil derfor ha større muligheter enn *Homo Oeconomicus* til å skape et stabilt, harmonisk samfunn, der folk tar vare på hverandre og stoler på hverandre. Men i det øyeblikket man passerer en grense der for mange tar seg til rette og utnytter andres tillit og innsatsvilje, vil hele harmonien kunne rakne, og det til gagns.

Assosiasjonen til velferdsstaten ligger snublende nær, så la meg trekke den et hakk lenger. Hvis mange er gjengjeldere, kan det utmerket godt tenkes at det finnes (minst) to potensielle

¹⁴ Fischbacher, U., og S. Gächter (2009): Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Good Experiments, kommer i *American Economic Review*.

¹⁵ Strengt tatt gir det neppe mening å snakke om preferansene til et land, men hvis majoriteten i et demokratisk land er gjengjeldere, vil de også kunne velge ledere som opptrer som gjengjeldere.

¹⁶ Se Sobel, J. (2005): "Interdependent Preferences and Reciprocity", *Journal of Economic Literature* 43, 392-436.

samfunnstilstander som begge, hvis den ene eller andre tilstanden først er oppstått, vil kunne være økonomisk og politisk stabile: Den minimalistiske staten, der enhver er sin egen lykkes smed og trykdesystemet er svært begrenset på grunn av faren for utstrakt misbruk; og velferdsstaten, der trykdeordningene er sjenerøse, men likevel misbrukes lite.¹⁷ Hvis alle var som *Homo Oeconomicus*, ville den siste tilstanden neppe kunne vært stabil: *Homo Oeconomicus* ville benyttet enhver anledning til å skulke jobben og snylte til seg offentlige penger, og systemet ville blitt både ineffektivt og svært dyrt. Hvis en stor del av befolkningen er gjengjeldere, derimot, fortøner ikke lenger velferdsstaten seg som noen indre motsigelse. Gjengjelderne kan tenkes å avstå fra å utnytte systemet, forutsatt at også andre avstår. For å kunne opprettholde velferdsstaten, blir det imidlertid overmåte viktig å holde gratispassasjerene i sjakk: Hvis de (muligens få) som skruppelløst snylter på systemet får utfolde seg fritt, kan de undergrave innsatsviljen til gjengjelderne, og slik undergrave stabiliteten i hele systemet.¹⁸

Gjensidighetspreferanser kan på denne måten skape gode eller dårlige sirkler når det gjelder vilje og evne til samarbeid om det felles beste. Også på den enkelte arbeidsplass kan slike mekanismer lede til enten godt eller dårlig samarbeidsklima. Visse studier indikerer at det kan være lønnsomt for arbeidsgivere å behandle arbeidstakerne bra, fordi mange da jobber hardere, selv om arbeidstakeren ikke vinner økonomisk på sin økte innsats.¹⁹ Hvis arbeidstakerne i tillegg betinger egen arbeidsinnsats på innsatsen til de andre i bedriften, vil de kunne påvirke og forsterke hverandres arbeidsmoral, enten den nå i utgangspunktet var god eller dårlig.²⁰

Normer dreier seg selvsagt ikke bare om gjensidighet og gjengjeldelse. Økonomer har også, innenfor rammene av formelle teoretiske modeller, analysert også mer eksplisitt etisk/moralsk motivasjon.²¹ Også det faktum at vi mennesker higer etter anerkjennelse og aksept fra andre, kan ha betydelige økonomiske og samfunnsmessige konsekvenser; og hvis det er mest anerkjennelse å få fra dem som opptrer omtrent som oss selv, vil dette gi en tendens til å løpe i flokk – noe som kan gi gode eller dårlige sirkler på liknende vis som beskrevet over.²² Professor Mari Rege ved Universitetet i Stavanger og jeg fant for eksempel at skjerpningen av røykeloven i 1988 kan ha endret normene for

¹⁷ Dette er et viktig forskningstema for *ESOP*, et av Norges forskningsråds sentre for fremragende forskning, som holder til på Økonomisk institutt ved UiO (<http://www.esop.uio.no/>).

¹⁸ For et beslektet argument, se Lindbäck, A. (1995): Welfare State Disincentives with Endogenous Habits and Norms, *Scandinavian Journal of Economics* 97(4), 477-494.

¹⁹ Se f.eks. A. Falk og M. Kosfeld (2006): The Hidden Cost of Control, *American Economic Review* 96 (5), 1611-1630.

²⁰ Se Brekke, K.A., K.E. Hauge, J.T. Lind, og K. Nyborg (2009): Playing with the Good Guys: A Public Good Game with Endogenous Group Formation, *Memorandum* 8/2009, Økonomisk institutt, UiO.

²¹ Et eksempel er Brekke, K. A., S. Kverndokk, og K. Nyborg (2003): An Economic Model of Moral Motivation, *Journal of Public Economics* 87 (9-10), 1967-1983; se også Tirole, J., og R. Benabou (2006): Incentives and Prosocial Behavior, *American Economic Review* 96(5), 1652-1678.

²² Rege, M. (2004): Social Norms and Private Provision of Public Goods, *Journal of Public Economic Theory*, 6 (1), p. 65-77.

innendørs røyking også på steder som ikke var omfattet av loven, som private hjem, nettopp via slike "sirkel"-mekanismer.²³

Selvkontroll

I atferdsøkonomien er det et tett samspill mellom teori og empiri. Hvis kartet ikke passer til terrenget, eller har hvite flekker, trenger kartet forbedring; men å lage bedre kart krever kjennskap til terrenget. Når eksperimenter systematisk forkaster prediksjoner fra etablerte teorier, stimuleres ny teoriutvikling, og når et forslag til ny teori foreligger, vil teorien typisk ha implikasjoner som igjen kan testes eksperimentelt. La oss derfor til slutt ta en titt på atferdsøkonomisk teoriutvikling, ved en enkel modell for *begrenset selvkontroll*.

For å gjøre det, må vi introdusere *tid* i analysen. *Homo Oeconomicus* er meget dyktig til å lage, og holde, langsiktige planer. I disse planene veier han nytte og kostnader i dag mot nytte og kostnader i framtiden. Jo mer utålmodig han er, jo mindre vekt vil han legge på framtiden; men han kjenner sin egen grad av utålmodighet, tar hensyn til den når han planlegger, og med mindre uforutsette hendelser oppstår, vil han også følge sine egne planer. Hans tenkning om fremtid versus nåtid er like planmessig, og følger omtrent samme prinsipper, som når banken beregner renter og rentesrenter på et fastrentelån.²⁴ Fristelse, viljessvakhet og indre konflikt er ukjente begreper for *Homo Oeconomicus*. Han vil derfor neppe forstå andres behov for å unngå å ha skåler med non-stop stående på kontorpuhlen, for å ha en absolutt deadline for å overhodet få skrevet sine artikler, eller for å begrense hvor mye penger og hvilke kredittkort man tar med ut på byen.

Kunnskap fra økonomiske eksperimenter så vel som biologi, psykologi og nevrologi tyder imidlertid på at både dyr²⁵ og mennesker²⁶ i praksis lider av (eller kanskje snarere, nyter) det som kan kalles *nåtidsskjevhet*: Vi legger mer vekt på øyeblikket, nået, enn fremtidige tidspunkt.²⁷ En nåtidsskjev person vil ha en tendens til å forandre planene etter hvert som tida går, og vil oppleve selvkontrollproblemer og indre konflikter: Vi synes vi bør slanke oss, men fristes over evne og utsetter det til i morgen, og i morgen utsetter vi det igjen. Selv hadde jeg vondt i en tann et helt år uten å ringe til tannlegen, og var overbevist om, helt til tannlegeinnkallingen lå i postkassa ganske av seg selv, at jeg skulle ringe til uka.

²³ Nyborg, K., og M. Rege (2003): On Social Norms: The Evolution of Considerate Smoking Behavior, *Journal of Economic Behaviour and Organization* 52, 323-340.

²⁴ Spesialservice for økonomer: Det jeg prøver å si her er "max $U = u(c^t) + \sum_{\tau=t+1}^{\infty} (\delta^{(\tau-t)} u(c^\tau))$ ", der c^t er konsum i periode t , δ er en konstant diskonteringsfaktor og $0 < \delta < 1$ ". (Hjalp det?)

²⁵ J. H. Kagel, R. C. Battalio, L. Green, *Economic Choice Theory: An Experimental Analysis of Animal Behavior*, Cambridge University Press, Cambridge, 1995.

²⁶ F.eks. Thaler, R.H. (1981): Some Empirical Evidence on Dynamic Inconsistency, *Economics Letters* 8, 201-207.

²⁷ Mer spesialservice for økonomer: Dette høres kanskje akkurat ut som vanlig eksponensiell diskontering, men er det ikke. Det kalles kvasi-hyperbolsk diskontering, fører til tidsinkonsistens, og ser sånn ut, der $0 < \beta \leq 1$, jf. ellers note 25: Max $U = u(c^t) + \beta \sum_{\tau=t+1}^{\infty} (\delta^{(\tau-t)} u(c^\tau))$. Se Laibson, D. (1997): Golden Eggs and Hyperbolic Discounting, *Quarterly Journal of Economics* 62, 433-477.

En mulig forklaring er at ønsket om øyeblikkelig nytelse (eller vegringen mot øyeblikkelig ubehag) og langsiktig planlegging later til å være styrt av ulike deler av hjernen.²⁸ Disse delene er ikke nødvendigvis samkjørt, og dermed kan indre konflikter nær sagt være fysisk tilstede i hodet (lyder kjent?): Fornuften vil la sjokoladen ligge i skapet, mens den mer primitive krypdyr-delen av hjernen krever å få den straks. Planleggeren i oss tar hensyn til både nåtid og framtid. Men fordi en beslutning alltid må skje i et nå, og krypdyrhjernen, som bare bryr seg om nå, dermed alltid er til stede når valget tas, vil krypdyret ha en tendens til å trekke oss nærmere sjokoladen i skapet enn det planleggeren ønsker. Herav, altså, nåtidsskjevheter.

Nå kan det jo hende at den nåtidsskjeve er klar over sin egen nåtidsskjevheter og tar hensyn til den, for eksempel ved å være å handle inn sjokolade slik at skapet er tomt når søtsuget melder seg, eller ved å avtale med tannlegen at han skal sende innkallinger automatisk. Spørsmålet er om markedet kan dekke etterspørselen etter slike "binde seg til masten"-tjenester. Svaret er av og til, spesielt hvis det er snakk om en positiv forpliktelse.²⁹ For eksempel tilbyr markedet kontrakter for bruk av helsestudio der man betaler mye på forhånd, men lite eller ingenting ved hver treningsøkt, en deal som subsidierer planleggeren i meg på bekostning av krypdyret. Markedet tilbyr også sparekontrakter med binde seg til masten-elementer, for eksempel ved at deler av gevinsten går tapt hvis jeg tar ut penger for ofte. Negative forpliktelser, som skal hindre meg i å falle for fristelser, kan derimot være atskillig verre å få til i markedet, slik historien nedenfor viser.

En tur på byen

Betrakt følgende eksempel: Periode 0 er en lørdag formiddag; jeg sitter hjemme og planlegger en tur på byen. Periode 1 er lørdag kveld, da jeg planlegger å være på bar. Periode 2 er søndag morgen.

I periode 0 skjer ingenting annet enn at jeg planlegger. Krypdyret i meg, som bare er opptatt av nået, har dermed ingen formening om hva jeg skal velge. Jeg planlegger å gå ut og ta maksimalt fire pils.³⁰ Da kan jeg ha en fin kveld, og også en akseptabel søndag morgen. Men så vet jeg, dessverre, at så snart jeg er på plass i baren, vil krypdyret melde seg. Søndag morgen vil da være senere, ikke nå, og jeg vil bli fristet til å drikke for mye. Dagen etter vil jeg angre.

Planleggeren i meg vet alt dette. Derfor bestemmer jeg meg for å ringe baren på forhånd og inngå en avtale med bartenderen: Jeg gir ham 100 kroner hvis han lover å ikke selge meg mer enn 4 øl, uavhengig av hva jeg forsøker å overbevise ham om lørdag kveld.

Så langt, så vel. Lykkelig og vel forlikte drar vi ut i verden, planlegger og krypdyr. Nøyaktig som forutsagt: Kvelden er ikke gammel før krypdyret vil forandre planene. Men nei: Bartenderen er en anstendig fyr, en avtale er en avtale, ikke prøv deg, jeg husker deg, du får ikke mer enn 4 pils.

²⁸ McClure, S.M., Laibson, D.I., Loewenstein, G. & Cohen, J.D. (2004). Separate neural systems value immediate and delayed monetary rewards, *Science* 304, 503-507.

²⁹ Köszeg, B. (2005): On the feasibility of market solutions to self-control problems, *Swedish Economic Policy Review* 12(2), 65-94.

³⁰ *Samtidens* konsulent hevder at fire pils er altfor mye for folk på min alder. Jeg innser at dette er sant, men selv planleggeren i meg ville nok insistert på fire, moderne ølglass er tross alt latterlig små.

Hva gjør krypdyret da? Enkelt. Krypdyret tar planleggeren i armen og drar henne med videre til neste bar, der ingen har laget noen avtale med bartenderen. Senere, når *nå* er søndag morgen og ikke lørdag kveld, angrer både krypdyret og planleggeren; det hjelper bare så lite.³¹ Jo flere konkurrerende tilbydere av øl som finnes på byen lørdag kveld, jo mindre sannsynlig er det at planleggeren i meg vil kunne kjøpe selvkontroll-hjelp i markedet.

Økonomer har vanligvis vært sterke tilhengere av avgifter ved det som kalles eksterne virkninger, altså at en handling får ukompenserte bivirkninger for andre (forurensning, for eksempel), fordi markedsprisene i slike tilfeller ikke fullt ut reflekterer de samfunnsmessige kostnadene ved handlingen. Avgifter i tilfeller uten eksterne virkninger, der den enkelte bare har seg selv å takke hvis det går galt, har hatt liten klangbunn blant økonomer; for ikke å snakke om reguleringer som åpningstidsrestriksjoner, forbud og så videre. Ut fra tradisjonelle økonomiske modeller har slike restriksjoner rett og slett ikke hatt noen god faglig begrunnelse. Men selvkontrollproblemer skaper faktisk en situasjon som likner veldig på en ekstern virkning. I stedet for en konflikt mellom f.eks. forurenser og forurensoffer, står konflikten nå mellom "jeg i dag" og "jeg i morgen". Hvis markedet ikke kan fremskaffe hjelpemidler til selvkontroll, er planleggeren i en viss forstand like hjelpeløs som forurensofferet, og begrunnelsen for f.eks. alkohol- og tobakksavgifter blir slett ikke så ulik begrunnelsen for miljøavgifter.

Miljøkonflikter bærer som regel i seg fordelingsmessige problemer: Hvem sine interesser skal tas mest hensyn til, forurenserens eller offerets? Etisk korrekte fasitsvar på dette finnes selvsagt ikke. Tilsvarende vil mine ulike jeg, henholdsvis hjemme, på bar, og dagen derpå, ha ulike prioriteringer, og kanskje er det ikke opplagt hvem av dem vi burde høre på. Later vi som om konflikten ikke finnes, er det imidlertid ingen tvil om hvem som tjener på det: krypdyret.

Hva nå?

Atferdsøkonomisk forskning gir nye perspektiver på mange viktige samfunnsspørsmål. Mens tradisjonelle økonomiske analyser har tendert til å betrakte fedme, rus og avhengighet som individuelle beslutninger, der myndighetene knapt har noen naturlig rolle, viser atferdsøkonomiske modeller at viljestyrkeproblemer ikke uten videre kan overlates til markedet. Mens eksistensen av den norske velferdsstaten nærmest har fremstått som en logisk inkonsistens i lys av tradisjonell økonomisk teori, gir atferdsøkonomien ny innsikt i hvorfor misbruk av velferdsordningene er begrenset – men også hvordan dette kan tenkes å endre seg.

Til tross for den store gjennomslagskraften atferdsøkonomisk tenkning har vist seg å ha i de økonomiske forskningsmiljøene, er forskningen fortsatt i en tidlig fase. Selv om det er påvist at arbeidsgivere kan tjene på å behandle sine ansatte bra, fordi de ansatte da jobber mer effektivt, vet vi ennå lite om når dette faktisk gjelder og når det *ikke* gjelder. Gjensidighetspreferanser *kan* bety at en ambisiøs, internasjonal klimaavtale faktisk er realistisk; men spørsmålet er ennå ikke grundig studert, og kanskje vil nærmere analyse vise at så enkelt er det dessverre ikke.

³¹ Utstrakt spesialservice for økonomer (se også note 30, og husk at $\beta < 1$): La U^t være preferansene som bestemmer faktiske valg i periode t . Da vil preferansene forandre seg over tid: $U^0 = \beta (u^1 + u^2)$ (hjemme; alle fremtidige perioder tillegges samme vekt), $U^1 = u^1 + \beta u^2$ (på bar; nå bryr jeg mindre om søndag morgen), $U^2 = \beta(u^1) + u^2$ (dagen derpå; nå synes kvelden før mindre viktig).

Det er lov å håpe at de nye atferdsøkonomiske modellene vil bidra til bedre politikk. Jeg tror likevel ikke en skal overdrive betydningen av dette: På mange av de områdene der *Homo Oeconomicus*-modellen gir sine mest hårreisende konklusjoner, er den, heldigvis, aldri blitt tatt helt på alvor av dem som utformer politikken. Vi *har* en ruspolitikk. Vi *har* en velferdsstat. Vi *har* offentlig pensjonssparing.

Imidlertid kan atferdsøkonomene få stor betydning som oversettere i den tverrfaglige kommunikasjonen innen samfunns- og atferdsfagene. Så vidt jeg kan se, deler disse fagene i dag verken virkelighetsforståelse, språk og begrepsapparat, eller metode.³² Atferdsøkonomien analyserer spørsmål som tradisjonelt særlig har vært sosiologers og psykologers domene, noe som har gjort det nødvendig for økonomer å lese sosiologisk og psykologisk litteratur – og dermed tilegne seg i det minste en viss forståelse for disse fagenes angrepsmåter. Atferdsøkonomene benytter seg av språk og metoder som andre økonomer også lett kan forstå. Dessuten presenterer de virkelighetsbeskrivelser som forhåpentlig fortøner seg noe mindre absurde for sosiologer og psykologer, og som derfor i større grad kan invitere til seriøs faglig debatt. Dersom atferdsøkonomien kan bidra til en mer helhetlig samfunnsvitenskap, vil det være en betydelig gevinst.

Homo Oeconomicus må uansett venne seg til at han ikke lenger, selv blant økonomene, er klassens eneste populære. Rivalene er i ferd med å bli mange, og de får økende oppmerksomhet. Det er bra. Kan hende besitter menneskene faktisk det potensialet som kreves for å kunne samarbeide om en bedre verden, ikke bare det som kreves for å ødelegge den. I så fall er det viktig at vi forstår, best mulig, hva som fremmer dette potensialet, og hva som bryter det ned. Det krever nitide og systematiske studier av menneskers atferd – som ikke begrenser seg til forestillingen om at mennesker kynisk og konsekvent grabber til seg alt de kan.

³² Se, for eksempel, flere artikler om dette i *Sosiolognytt* 3/2009.