

Nudge – Et skub i den rigtige retning?

**Felix Sebastian Døssing, Ph.d.-stipendiat, Københavns Universitet,
Økonomiske Institut, fsd@econ.ku.dk**

**David Dreyer Lassen, Professor, Københavns Universitet,
Økonomisk Institut, david.dreyer.lassen@econ.ku.dk**

**Daniel Gerszon Mahler, Ph.d.-stipendiat, Københavns Universitet,
Økonomiske Institut, daniel.mahler@econ.ku.dk**

Interventioner baseret på adfærdsøkonomisk viden udgør i stigende grad et populært værktøj i den offentlige forvaltning i Danmark såvel som på verdensplan. Vi diskuterer, hvorledes man kan vurdere, hvorvidt disse tiltag har en positiv eller negativ effekt på velfærden i samfundet.

Velfærdsforringende interventioner

Peter er enlig far til to børn. Han har en indtægt, der ligger lidt under den gennemsnitlige i Danmark og havde – indtil han fik nyt job – valgt at spare 1.500 kr. op til pension om måneden. Efter at han har skiftet job, er han imidlertid blevet tilmeldt en automatisk pensionsordning, hvor han indbetaler 15 pct. af sin indkomst. Som enlig far er der rigeligt med udgifter, og de automatisk indbetalte penge er således mere værd for Peter nu end i alderdommen. Han kunne naturligvis blot framelde sig den automatiske ordning, men i en travl hverdag er det ikke sikkert, han får taget sig sammen til at få det gjort, eller måske i det hele taget ikke opfatter, at ordningen faktisk kan frameldes. I dette tilfælde er det lykkedes at få Peter til at spare mere op til pension, men samtidig er hans situation blevet forværret. Vi kan konkludere, at interventionen – automatisk tilmelding af pensionsopsparing – har haft en negativ velfærdseffekt for Peter.

Ovenstående er et eksempel på en intervention, der er baseret på adfærdsøkonomisk teori, som i stigende grad anvendes som policy-instrument. Adfærdsøkonomi er en disciplin, der beskæftiger sig med betydningen af psykologiske aspekter for beslutningstagning og disses betydning for økonomisk teori og praksis. Mere overordnet kan adfærdsøkonomi imidlertid også ses som en insisteren på at forholde sig til, hvorledes folk faktisk handler i økonomiske sammenhænge, og at tilpasse økonomiske modeller til den observerede virkelighed. Adfærdsøkonomien har med stor succes dokumenteret økonomisk relevante fænomener såsom tabsaversion (Kahneman og Tversky, 1979), tidsinkonsistente præferencer (Loewenstein og Prelec, 1992), begrænset opmærksomhed og regnekraft (Simon, 1957) samt sociale præferencer (Charness og Rabin, 2002). På nuværende tidspunkt er den – i overbevisende grad – lykkedes med at kvalificere disse fund og i en vis grad at påvise deres betydning uden for laboratoriet (se f.eks. Camerer o.a., 1997; Gneezy og List, 2006). Som konsekvens heraf har der gennem de sidste 10-20 år været stor fokus på, hvorledes adfærdsøkonomiske fejl og bias kan håndteres i en policy-sammenhæng. Dette fokus skyldes især udgivelsen af bogen Nudge, hvor Thaler og Sunstein (2008) viste,

hvorledes interventioner baseret på adfærdsøkonomisk viden kan anvendes til at intervenere i folks adfærd. I tråd med disse vil vi referere til interventioner, der er adfærdsøkonomisk motiveret, som »nudges«.

I denne artikel vil vi diskutere et grundlæggende spørgsmål, der knytter sig til adfærdsøkonomiske interventioner: Hvorledes vurderer vi, hvorvidt en given intervention har en positiv effekt på folks velfærd? Det er (naturligvis) et helt fundamentalt spørgsmål: Hvis en intervention ikke øger velfærd, bør man ikke spille tid og penge på den. Ikke desto mindre er det ganske almindeligt, at man i litteraturen blot forholder sig til en interventions effekt på en given målvariabel (som f.eks. opsparing eller energibesparelser) og implicit antager, at mere af en sådan variabel altid er godt for den enkelte. Således forholder man sig ikke eksplicit til velfærdseffekter. Vi vil diskutere de problemer, der knytter sig til en vurdering af en interventions velfærdseffekt, og gennemgå en række nyere forskningsartikler, der på forskellig vis fremhæver, at dette spørgsmål kan være mere komplekst, end det kan forekomme ved første øjekast.

Det velfærdsøkonomiske udgangspunkt

Velfærdsøkonomi beskæftiger sig med spørgsmålet om, hvilke politiske valg der maksimerer individers velfærd. Begrebet velfærd anvendes her i sin klassiske betydning som synonym for nytte fremfor i sin ofte anvendte moderne betydning, hvor det typisk udgør et synonym for goder og services leveret af det offentlige. Vi tager udgangspunkt i et velfærdsøkonomisk perspektiv i denne artikel: *Samfundets* velfærd forstås således som en aggregering af individers velfærd. Der er i velfærdsøkonomien et grundlæggende ønske om at frigøre sin analyse fra ideologiske hensyn og blot tage udgangspunkt i individers præferencer, som Mas-Colell o.a. (1995) formulerer det: »Det er et traditionelt og fast princip inden for velfærdsøkonomien, at politisk beslutningstagning ikke bør være *paternali-*

stisk. Som minimum betyder dette, at alternativer, der ikke kan skelnes med udgangspunkt i agents præferencer, heller ikke bør skelnes af beslutningstageren« (p. 818; oversat af forfatter). Der kan naturligvis være andet end velfærd, der er værdifuldt for et samfund – såsom lighed eller demokratisk deltagelse. Her vil vi imidlertid fokusere på individers velfærd, da det ofte er dette, adfærdsøkonomiske interventioner har til hensigt at forbedre. Såfremt andre samfundsværdier påvirker individers velfærd – hvis mere lighed eller demokratisk deltagelse f.eks. øger individers livskvalitet – indgår disse værdier indirekte i velfærdsbetragtninger. Selvom vi her definerer velfærd som individers velfærd, kan det være nødvendigt med en yderligere præcisering af, hvad der menes med dette. Vi leverer en sådan præcisering i artiklens fjerde afsnit.

Potentialet ved at nudge

Adfærdsøkonomien kan siges at blive politikrelevant, når udgangspunktet for folks valg (deres valg-funktion) og deres nytte (nytte-funktionen) ikke er sammenfaldende (Kahneman o.a., 1997). Det kan for eksempel ske, hvis jeg vælger, hvilken bil jeg vil købe, alene ud fra pris og udseende, mens andre faktorer, såsom benzinforsbrug ender med at være afgørende for min nytte. Vi kan kalde forskellen mellem den nytte, jeg får ud af mit faktiske valg, og den nytte, jeg ville få ud af mit optimale valg, for et bias eller en »internalitet«. Ordet internalitet har til formål at pointere ligheden til »eksternalitet«, som er et mere anvendt økonomisk begreb. Dette kan forstås således, at mit nuværende jeks handling udgør en eksternalitet, idet den mindsker mit fremtidig jeks nytte. Accepterer man, at valg og nytte ikke altid er perfekt sammenfaldende, betyder det, at ikke alle mennesker kan antages til enhver tid at handle optimalt, og der vil således være potentiale for at øge velfærd ved at intervenere i folks adfærd. Dette kan enten være via mere klassiske instrumenter, såsom beskatning og subsidiering eller via nudges.

Den mest almindelige tilgang inden for adfærdsøkonomien er at måle effekten af en intervention på, hvad der a priori antages at være værd at fremme. Det kan f.eks. være pensionsopsparing, velgørende donationer, optimering af investeringer, sundere levevis osv. Adfærdsøkonomi kan være nyttigt til at ændre folks forbrug af et bestemt gode ved at forbedre vores forståelse af virkningen af klassiske instrumenter, såsom subsidier og beskatning og ved at stille nye og ofte mere omkostningseffektive instrumenter til rådighed. Her udgør forsøg på at øge pensionsopsparing et godt eksempel. Manglende opsparing til alderdommen kan dels udgøre en internalitet, idet individer vil ende med at fortryde deres valg, når de bliver ældre, og dels udgøre en eksternalitet, eftersom fællesskabet vil have vanskeligt ved at binde sig til ikke at hjælpe personer, der som konsekvens af manglende opsparing får brug for hjælp (det såkaldte *Samaritanerens dilemma*). Af disse årsager er der generelt en bekymring om manglende pensionsopsparing (Poterba, 2014) og et politisk ønske om at fremme opsparing.

Den klassiske tilgang til dette problem er at forsøge at fremme opsparing via subsidier. Set fra et adfærdsøkonomisk perspektiv bør man imidlertid forvente, at dette instrument kun er virksomt, såfremt folk både er opmærksomme på subsidiet og formår at ræsonnere sig frem til den passende reaktion. Derfor er det relevant at overveje, hvorvidt nudges vil være en mere effektiv måde at øge opsparing på.

Dette illustreres af Chetty o.a. (2014), der kigger på effekten af ændringen i subsidiering af pension på danskernes opsparingsadfærd samt effekten af automatisk tilmelding. I 1999 blev fradraget på opsparing af kapitalpension reduceret fra 59 øre per krone til 45 øre per krone for personer med høj indkomst. Ud fra klassisk efterspørgselsteori bør den marginale ændring i subsidiet føre til

en ændring i opsparing, der er forskellig fra nul. I modsætning hertil finder Chetty o.a., at størstedelen af dem, som blev påvirket af ændringen, ikke ændrer deres opsparing som konsekvens af prisændringen. Forfatterne konkluderer, at omkring 85 pct. af befolkningen udviser »passiv« opsparingsadfærd, hvilket vil sige, at de ændrer deres opsparing, hvis deres betaling ændres automatisk, mens de ikke reagerer på prisændringer. Af de »aktive«, der ændrer deres opsparing som konsekvens af prisændringen, vælger de fleste at flytte opsparing over til andre former for opsparing, således at sænkningen af subsidiet samlet set har en meget lille effekt på opsparingen. Deres analyse viser i øvrigt, at de, der reagerer på prisændringen, lader til at være de mest økonomisk sofistikerede og således ikke er den gruppe, som er i risiko for at ende uden tilstrækkelig opsparing.

Ønsker man at ændre opsparingsadfærden blandt de passive, er et oplagt middel en ændring i den automatiske indbetaling. Det er et langt billigere instrument og har potentiale for at være langt mere effektivt. Ifølge den klassiske model tilpasser individer deres opsparing til deres ønskede niveau, og automatisk tilmelding bør således ikke have nogen effekt, idet man blot kan modvirke den automatiske indbetaling ved at sænke sin frivillige opsparing; men Chetty o.a. finder, at folks totale opsparing i gennemsnit stiger med 80 øre, når de skifter job til en virksomhed med 1 kroner større automatisk indbetaling, selvom de har mulighed for at modvirke dette ved at sænke deres private opsparing. Her ses altså et eksempel, hvor et nudge er langt billigere og langt mere effektivt end det klassiske instrument. Det er ét blandt mange eksempler på, hvorledes nudges kan levere omkostningseffektive ændringer i folks adfærd.

Negative konsekvenser ved at nudge

På trods af at nudges kan være et effektivt redskab til at øge samfundets velfærd ved

at ændre specifik adfærd såsom opsparing, motion osv., er der en række problemer, der knytter sig til brugen af nudges; vi ser på nogle af disse problemer i det følgende.

Hjælper interventionen målgruppen?

Ser vi alene på adfærdsændringens betydning for de personer, som interventionen er rettet mod, er det først og fremmest vanskeligt at afgøre, hvorvidt der samlet set sker en forøgelse af velfærd for disse personer. Grundlæggende er problemet, at vi reelt ikke ved, hvorvidt de påvirkede ønsker mere af det specifikke gode. Selvom det er relativt uskyldigt at antage, at alle ville være bedre stillet, hvis de *alt andet lige* havde bedre sundhed, mere opsparing osv., sker adfærdsinterventioner ofte på bekostning af et andet værdifuldt gode. Hvis vi ved brug af nudges f.eks. får folk til at spare mere op, sker dette som regel på bekostning af deres nuværende forbrug. Hvis vi ikke har nogen ide om, hvorledes vi opgør og sammenligner velfærdseffekten af f.eks. øget opsparing kontra øget forbrug, er det vanskeligt at afgøre, hvorvidt ens intervention gør skade eller gavn. I indledningen til denne artikel præsenterede vi et eksempel på en person, der reagerer på adfærdsinterventionen, men som netop havde mere gavn af nuværende kontra fremtidigt forbrug.

Selv hvis det vurderes, at interventionen har en positiv effekt på målgruppen, er det værd at være opmærksom på, at nudges ofte blot løser en lille andel af de problemer, som suboptimal beslutningstagning medfører. Hvor automatisk tilmelding af pensionsopsparing f.eks. har vist sit værd som instrument til forøgelse af opsparing, er det langt fra en fuldkommen løsning på det problem, at størstedelen af befolkningen ikke forholder sig aktivt til egen opsparing. Som pointeret af Stott og Nottingham (2017) vil automatisk tilmelding næppe medføre, at folk rammer deres optimale opsparingsniveau og deres ønskede risiko-afkast-profil. Dette understreges af en stor heterogenitet i investeringsstrategierne

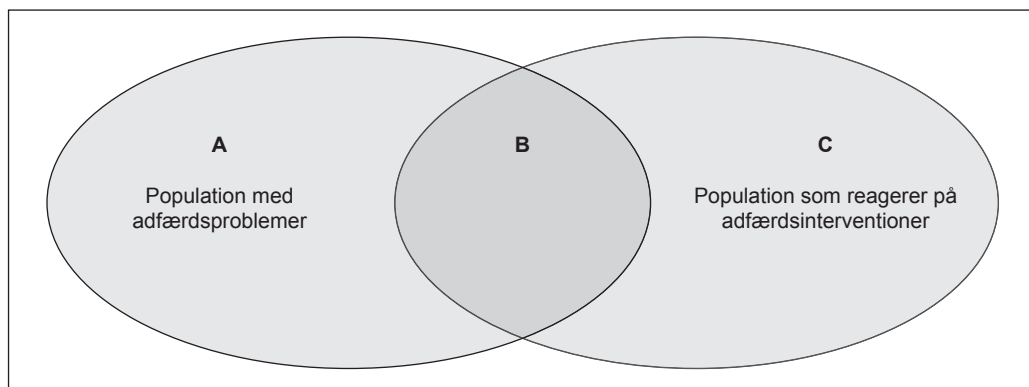
hos de pensionsfonde, som folk tilmeldes (i hvert fald i en britisk kontekst). Herudover er der risiko for, at passiv investering forøger den principal-agent problematik, der altid er knyttet til fondsinvestering, hvilket kan medføre misligholdelse af folks investering. Stott og Nottingham vurderer, at denne problematik kan medføre et enormt tab af opsparing for den enkelte, og anbefaler, at man frem for blot at anvende automatisk tilmelding ligeledes gør en indsats for (1) at gøre folk opmærksom på det tab, som deres manglende engagement medfører; (2) at gøre opmærksom på, at standardtilmeldte fonde ikke skal forstås som en anbefaling; og (3) at investere i uddannelse for at øge finansiell forståelse. Det er klart, at det at sætte sig ind i investeringen af ens pension for de fleste også vil involvere en betydelig alternativomkostning. Dette komplicerer blot tingene yderligere, da det er vanskeligt at vurdere, hvornår denne alternativomkostning opvejer fordelene ved aktiv forvaltning. Ikke desto mindre er uddannelse for at øge finansiell forståelse et godt bud på, hvorledes man kan sænke alternativomkostningerne og samtidig øge den aktive forvaltning.

Har interventionen konsekvenser for folk uden for målgruppen?

Det vil ofte blot være en delmængde af dem, der påvirkes af en intervention, der faktisk har problemer med selv at handle optimalt. Ved stort set alle adfærdsinterventioner er der nogle, der 'rationelt' forbruger det specifikke gode på det nuværende niveau: Der er nogle mennesker, der har afvist sundhedsomkostningerne ved rygning og stadig finder, at det er optimalt at ryge, nogle mennesker, der ikke sparer op af deres indkomst grundet deres økonomiske situation eller forventede levetid, osv.

Man kan overordnet set skelne mellem tre relevante grupper, jf. figur 1. Der er en gruppe med adfærdsproblemer, som ikke lader sig påvirke af interventionen (gruppe A), og for

Figur 1: Reaktionen på adfærsinterventioner



hvilken velfærdseffekten af interventionen derfor vil være lig nul. Dette kunne eksempelvis være personer, som sparer for lidt op til deres pension, men som fortsætter med dette uafhængigt af opsparings subsidier eller automatisk indbetaling. Dernæst er der en gruppe af mennesker med adfærdsproblemer, hvis adfærd korrigeres af interventionen (gruppe B). Velfærdseffekten af interventionen for denne gruppe er positiv, såfremt adfærdsændringen ikke er for meget større end personens bias. Endelig er der den gruppe af mennesker, der ikke har adfærdsproblemer i det specifikke domæne, men som stadig bliver påvirket af interventionen (gruppe C). Denne gruppe handlede allerede optimalt før, hvorfor en ændring af deres adfærd formentlig fører til et fald i deres velfærd. Dette kunne være folk, for hvem det ikke er optimalt at spare mere op til pension, men som ikke desto mindre reagerer på en indførsel af automatisk indbetaling grundet passivitet. Det er vanskeligt at vurdere, hvorvidt den positive effekt på gruppe B overstiger den negative effekt på gruppe C. For at vurdere dette er man for det første nødt til at vælge, hvorledes man har tænkt sig at måle ændringer i velfærd for den enkelte, og herudover kræver det, at man vælger en metode til at sammenstille velfærdsændringer på tværs af individer.

At sammenligne velfærd på tværs af individer udgør et klassisk problem i økonomisk teori. Det gør det vanskeligt at vurdere samfundets samlede velfærd samt politiske tiltags betydning for den samlede velfærd. For at forstå hvorfor dette er tilfældet, kan vi jf. tabel 1 betragte to arbitrære personers præferencer:

Tabel 1: Sammenligning af præferencer

Person A	Person B
⋮	⋮
b	β
⋮	⋮
c	γ
⋮	⋮
d	δ
⋮	⋮

De små bogstaver angiver forskellige forbrugsbundter, hvor et bundt højere oppe på listen foretrækkes over et bundt længere nede på listen. Person A foretrækker således bundt b over bundt c og så videre. Symbolet »:« refererer til, at der er uendeligt mange bundter placeret mellem bundt b og c , og ligeledes mellem de andre forbrugsbundter. Der kan placeres uendeligt mange bundter imellem to vilkårlige bundter, da man altid kan tilføje marginalt ekstra indkomst (eller et andet kon-

tinuert gode som en person foretrækker mere af) til et bundt.

Lad os betragte et udgangspunkt, hvor Person A forbruger bundt d , mens person B forbruger bundt γ . Omend man fristes til, at konkludere, at B er bedre stillet end A, da B er 'tættere på toppen', holder dette ikke, da de begge har uendeligt mange muligheder, de foretrækker frem for deres nuværende bundt. Forestil dig nu, at vi vedtager et politisk tiltag, som får A til at forbruge bundt b , mens B kommer til at forbruge bundt δ . Vi kan heraf konkludere, at A er blevet bedre stillet og B er blevet værre stillet. Det er dog ikke muligt at afgøre om samfundet som helhed er bedre stillet, da vi ikke kan sammenligne størrelsen af disse ændringer – de to personer er hoppet uendeligt mange skridt i forskellig retning. Derfor vurderer økonomer ofte politiske tiltag ved at se, om de gør visse personer bedre stillet, uden at nogen bliver værre stillet. I så fald kaldes det politiske tiltag for en *paretoforbedring*, hvilket i daglig tale går under navnet *efficiens*.²

Kan et bias være velfærdsforøgende?

Selv hvis folk uden for målgruppen ikke reagerer på interventionen, og population C således ikke eksisterer, er velfærdseffekten af at korrigere et bias ikke altid utvetydig. Faktisk kan bias være velfærdsforøgende, og korrektion af disse risikerer at have negative konsekvenser. Et eksempel på dette ses hos Rees-Jones og Taubinsky (2016), der undersøger de heuristikker, som folk anvender til at beregne deres indkomstskat. Disse heuristikker fører generelt til en undervurdering af skatten forbundet med øget indkomst og overvurdering af skatten forbundet med sænket indkomst. Med udgangspunkt i dette viser de, at nudges, der forbedrer folks forståelse af beskatning, kan få folk til at arbejde mindre. I dette tilfælde vil en øget forståelse af beskatningen på den ene side føre til optimering for den enkelte, men samtidig føre til færre skatteindtægter og en forøgelse af det dødvægstab, der er forbundet med beskatningen.

Et andet eksempel på en adfærdsintervention, der ved at korrigere et bias medfører et velfærdstab, fremgår af Chetty o.a. (2009). Forfatterne viser, at den økonomiske konsekvens af *sales tax* afhænger af, om beskatningen anmærkes på varens prisskilt eller tillægges ved selve købet. Ud fra dette fund konstruerer de en model, der illustrerer dilemmaet forbundet med at tydeliggøre beskatningen på prisskiltet: Skal vi opgøre velfærdseffekten af manglende opmærksomhed, afhænger vores konklusion af det perspektiv, vi tager. Ser vi alene på forbrugerens nytte, medfører manglende opmærksomhed et tab. Til gengæld har regeringen en forøgelse i indsamlet skat. Ser vi på den samlede effekt, medfører manglende opmærksomhed, at dødvægstabet forbundet med beskatningen vil falde, hvilket vil sige, at den samlede velfærd stiger som konsekvens af manglende opmærksomhed og at forsøg på at »de-biase« eller kalibrere forbrugeren vil medføre et velfærdstab, såfremt det lykkes. Det er værd at bemærke, at selvom forbrugeren lider et tab grundet sin manglende opmærksomhed, så kan regeringen i princippet kompensere vedkommende for dette tab og stadig have en større skatteindtægt (ibid.).

Analogt med dette har en række artikler vist, hvorledes nudges, der hjælper den enkelte til at optimere sit valg af sundhedsforsikring, kan have store negative effekter, idet de forstærker det problem, der betegnes som *adverse selection* (Handel, 2013; Handel og Kolstad, 2015): Ved at hjælpe folk med at overkomme den inerti, der er forbundet med optimalt valg af forsikring, medfører det, at personer med bedre helbred vælger mindre dækning, hvilket gør det dyrere for personer med dårligt helbred. Dette udgør en (klassisk) markedsfejl, som nudging risikerer at forstærke.

Det faktum, at fejl i folks beslutningstagning faktisk kan være velfærdsforøgende, udgør naturligvis en etisk hveserede. Hvis vi f.eks.

finder, at det er velfærdsoptimerende at fratække moms fra folks løn frem for direkte ved køb (hvis dette var muligt) og dermed lade dem forbruge mere, end hvad der er individuelt optimalt, er det så etisk forsvarligt? Dette er et godt eksempel på, hvorledes adfærdsøkonomien *nødvendiggør* etisk og ideologisk stillingtagen. Ser vi det som etisk forsvarligt at løse samfundsmæssige problemer via fortællinger og fremprovokering af misforståelser? Hvilken vægt placerer vi på individets autonomi eller på individets ret til at træffe dårlige valg? Det sidste spørgsmål vil vi diskutere yderligere i artiklens femte afsnit.

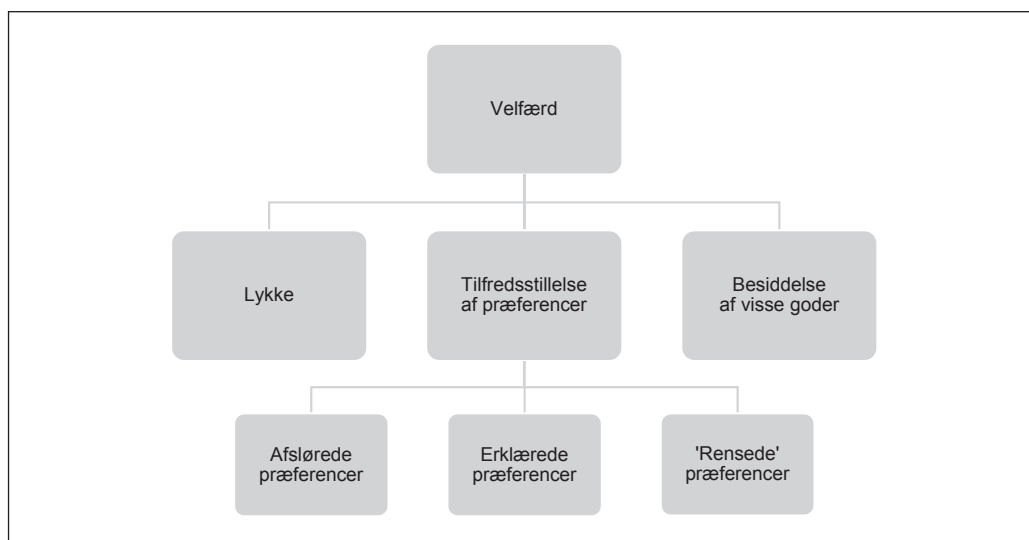
Hvad er velfærd?

Som tidligere nævnt har vi som udgangspunkt, at velfærd består i individers velfærd, og at *samfundets* velfærd må forstås som en aggregering af individernes velfærd alene. Ønsker vi at afgøre velfærdseffekten af en given intervention, kan det imidlertid være nødvendigt med en yderligere forståelse af, hvad der udgør individers velfærd og hvorledes vi kan måle dette. I filosofien skelnes der, jf. figur 2, typisk mellem tre forskellige

opfattelser af, hvad velfærd er: (1) besiddelse af goder, (2) tilfredsstillelse af præferencer og (3) lykke (Parfit, 1984; Griffin, 1986; se figur 2).

Defineres velfærd som *besiddelsen af visse goder*, antages det, at der er en liste af goder, der tilsammen udgør ens velfærd. Disse goder kunne være indkomst, helbred, uddannelse osv. Det er hovedsageligt denne velfærdsopfattelse, som adfærdsøkonomi (implicit) har anvendt. Ofte måles effekten af adfærdsinterventioner nemlig på forbruget af et gode, der a priori antages at være værd at fremme. Typisk identificeres et gode, som man mener underforbruges, og en eller flere interventioner, som man mener vil påvirke forbruget af dette gode. Herefter evalueres (velfærds)effekten ud fra, hvorvidt interventionen lykkes med at ændre forbruget. 'Goder' kan forstås bredt – det kan f.eks. være pensionsopsparing, optimering af investering, sundere levevis osv. Som tidligere nævnt er det typisk denne forståelse af velfærd, man tager udgangspunkt i, når man evaluerer adfærdsøkonomiske interventioner.

Figur 2: Opfattelser af velfærd



Der er en række problematikker, der knytter sig til denne forståelse af velfærd. Som tidligere nævnt baserer velfærdsøkonomien sig på et ønske om at undgå paternalisme. Det er dog vanskeligt at undslippe en vis paternalisme, hvis vi tillader os at definere på folks vegne, hvilke goder de har godt af at forbruge. Mere praktisk er det langt fra entydigt, hvad der skal indgå på listen af goder, der giver velfærd. Selvom det er relativt uskyldigt at antage, at alle ville være bedre stillet, hvis de *alt andet lige* havde bedre sundhed, mere opsparing osv., opstår der et nyt problem, i og med at adfærdsinterventioner ofte sker på bekostning af et andet værdifuldt gode. Denne problematik har vi tidligere nævnt, bl.a. eksemplificeret ved afvejningen mellem nytten forbundet med nuværende forbrug og nytten forbundet med senere forbrug som konsekvens af opsparing. Hvis vi ikke har nogen idé om, hvorledes vi opgør og sammenligner velfærdseffekten af øget opsparing kontra øget forbrug, er det vanskeligt at afgøre, hvorvidt ens intervention gør skade eller gavn. For at vurdere om tiltaget samlet har en positiv effekt, kræver det, at man kan veje disse to ændringer op mod hinanden, hvilket forudsætter både detaljeret information og nogle potentielt ret arbitrære og/eller subjektive valg fra beslutningstagernes side.

En umiddelbar løsning på dette problem er at lade folk selv bestemme, hvor vigtige goder er for deres velfærd. Det er netop denne tilgang, som økonomer traditionelt har anvendt, når de har taget udgangspunkt i folks præferencer. Velfærd defineret som *tilfredsstillelsen af præferencer* bygger på idéen om, at mennesker har præferencerrelation, ud fra hvilken de kan rangere forskellige udfald ud fra, hvad der er bedst for dem. Jo højere op på præferencestigen man kommer, jo mere velfærd har man. Ønsker man at måle folks præferencer, kan man blandt andet gøre dette ved at observere deres adfærd og beslutninger. Hvis en person kan vælge mellem A og B og vælger A, kan det antages, at denne per-

son foretrækker A frem for B. Denne tilgang kaldes for *afslørede præferencer*, idet det antages, at folk afslører deres præferencer via de valg, de træffer. Velfærdsøkonomi har traditionelt beskæftiget sig med tilfredsstillelse af præferencer og har anvendt afslørede præferencer som velfærdskoncept (Mas-Colell o.a., 1995). Når økonomer eksempelvis taler om 'optimal' beskatning eller 'velfærdsmaximerende' politikker, bygger dette på studier, hvor individers adfærd og beslutninger observeres. En af hovedårsagerne til, at økonomer anvender dette velfærdsbegreb, er, at det bliver anset for ikke at være paternalistisk: Individer bestemmer selv, hvordan de afvejer forskellige forbrugsbundter, og således hvad der skal til, for at for at de opnår velfærd.

Et problem med at anvende afslørede præferencer er imidlertid, at adfærdsøkonomer og psykologer som bekendt har vist, at individer i mange tilfælde ikke træffer optimale valg grundet manglende information, manglende viljestyrke eller andre biases. Hvis folk ikke træffer optimale valg, er der ikke belæg for at antage, at afslørede præferencer reflekterer velfærd. Hvis en person f.eks. vælger ikke at have en pensionsopsparing, kan det ikke længere antages, at vedkommende foretrækker ingen opsparing frem for opsparing, idet valget kan skyldes undervurdering af fremtidig værdi eller uopmærksomhed. Det er implicit en af årsagerne til, at adfærdsøkonomi kun sjældent anvender afslørede præferencer til at vurdere politiske tiltag. I stedet for at skifte velfærdskonceptet ud er det dog muligt at revidere præferenceforståelsen, således at det kan anvendes til at vurdere konsekvenserne af adfærdsinterventioner. Dette kan blandt andet gøres ved at anvende *erklærede præferencer*. Her måles præferencer, ved at folk erklærer, om de finder gode A bedre end gode B. Dette kan eksempelvis måles ved hjælp af surveys. Hvis folk er bevidste om deres adfærdsproblemer, vil dette reflekteres i erklærede præferencer, hvorfor de kan formå at løse nogle af problemerne med afslørede

præferencer. Det er dog usandsynligt, at folk er fuldt bevidste om deres adfærdsproblemer og samtidig besidder den nødvendige viden til altid at vurdere hvilket gode, der er bedst for dem. Begge disse problemer kan løses bedre ved at tage udgangspunkt i såkaldt 'rensede' præferencer. Ved brug af rensede præferencer accepterer man, at mange ikke nødvendigvis træffer optimale valg, og forsøger at korrigere ('rense') individets præferencer herefter. Når folk eksempelvis undervurderer værdi i fremtiden, hvilket vil sige, at de diskonterer fremtiden mere, end hvad de selv finder optimalt set over et længere perspektiv, kan vi forsøge at finde frem til personens »langsigtede præferencer«. Dette kan vi f.eks. gøre via eksperimenter. O'Donoghue og Rabin (2003) følger denne strategi til at vurdere den optimale beskatning på 'synd-goder' såsom cigaretter og sodavand. En ulempe ved denne metode er, at det er en forudsætning, at de relevante adfærdsbias kan formuleres ret præcist. Man skal således vide, hvordan folk ville agere, hvis de var fuldt informerede og fuldt rationelle. Vides dette ikke, kan denne metode ikke anvendes.

Grundet problematikken med at anvende præferencer, når folks valg lider af bias, er nogle økonomer begyndt at anvende *lykke* som velfærdskoncept. I praksis bliver lykke ofte målt ved at spørge folk, hvor tilfredse de er med deres liv på en veldefineret skala. Hvis vi måler individets lykkeniveau før og efter en adfærdsintervention, kan det vurderes, om denne har en positiv eller negativ effekt på det gennemsnitlige lykkeniveau. Gruber og Mullainathan (2005) tager udgangspunkt i denne fremgangsmåde til at vurdere, om cigaretskatter har en positiv effekt på rygeres velfærd, hvilket de konkluderer er tilfældet. Det er værd at nævne, at det at tage udgangspunkt i målt lykke ikke ophæver problemet med sammenligning på tværs af individer. Det kræver nemlig, at folk forstår de anvendte lykkeskaler på samme måde. Er dette ikke tilfældet, er det vanskeligt at vide, hvem

der er bedst stillet. Sagt med andre ord, hvis person A's selvrapporterede lykke stiger med 2, mens person B's falder med 1, er det vanskeligt at konkludere noget om samfundets samlede lykke, idet vi ikke kan vide, om den subjektive ændring, der bliver til en ændring i rapportering på 2 for A, faktisk er større end den subjektive ændring, der resulterer i en ændring i rapportering på 1 for B. Såfremt det vurderes, at subjektive lykkeniveauer er troværdige, formår denne metode dog at afveje både de positive og potentielt negative velfærdseffekter af adfærdsinterventioner.

Der er således en række andre metoder til at måle velfærd end den sædvanlige tilgang i adfærdsinterventioner. Om end alle disse metoder har deres svagheder, kan det være værd at forsøge sig med forskellige tilgange. Hvis flere velfærdsmål peger i den samme retning, vil konklusionen alt andet lige være mere robust.

Værdien af frihed

Indtil nu har vi diskuteret adfærdsøkonomiens betydning for vores forståelse af velfærd. Der er imidlertid også bredere politiske konsekvenser, som vi her blot vil forsøge at ridse op helt kort. Økonomisk teori har historisk været tæt forbundet til liberalismen exemplificeret ved tænkere som Smith, Mill og Locke og senere Hayek og Friedmann. På belejlig vis var den neoklassiske økonomiske teori i overensstemmelse med principper som autonomi og personlig frihed (Saint-Paul, 2011): Antager man, at alle mennesker handler rationelt og har komplet information, vil minimal intervention være optimal. Velfærdsøkonomiens første hovedsætning viste under visse antagelser, at frihed (det friemarked) fører til en Pareto-optimal tilstand. Dette vil sige, at ingen kan stilles bedre, uden at andre derved stilles værre. Pareto-optimale tilstande kan forstås som et minimumskrav for, at et udfald giver tilstrækkelig velfærd. Det at frihed er blevet set som effektiv, bør imidlertid ikke gøre os blinde for, at der er

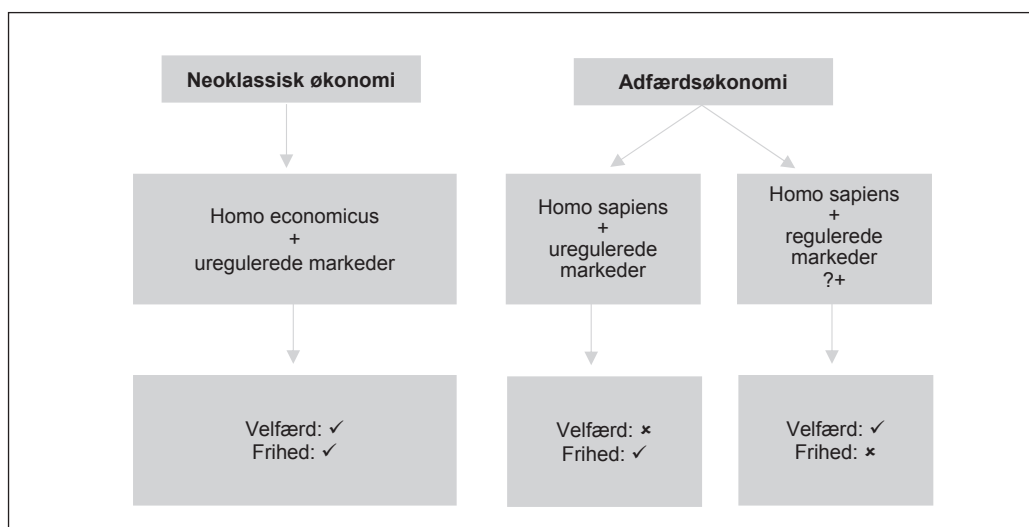
tale om to adskilte forsvar for autonomi: Man kan ønske autonomi, fordi man tror på, at det maksimerer nytte, eller man kan ønske autonomi, fordi man mener, at autonomi har værdi i sig selv.

Om end der i klassisk økonomi er tilfælde på flækker i denne harmoni – særligt når vi har med eksternaliteter eller offentlige goder at gøre – repræsenterer adfærdsøkonomien måske det hidtil største teoretiske brud med harmonien mellem frihed og velfærd. Godtager man pointen om, at folk langtfra altid handler optimalt, så bliver det nødvendigt at veje velfærd op over for principper som frihed og autonomi (figur 4). I praksis betyder det adfærdsøkonomiske perspektiv, at interventioner, der begrænser folks frihed, godt vil kunne øge deres velfærd, og at der således vil være et trade-off mellem velfærd og frihed. Selv nudges, hvilke per definition strengt set ikke udelukker muligheder, påvirker stadig folks autonomi forstået som deres frihed fra ekstern indflydelse. Det bliver altså også afgørende, hvorvidt vi definerer frihed som det

tekniske sæt af valgmuligheder eller som frihed fra påvirkning og manipulation.

Hvis frihed har en separat værdi, er det endnu vanskeligere at vurdere, om adfærdsinterventioner samlet set har en positiv effekt. Nogle velfærdsbegreber er brede nok til at kunne inkludere frihed i dets omfang. Anvendes eksempelvis besiddelsen af visse goder, kan man tilføje frihed til listen af goder, der har betydning. Anvendes præferencer som velfærdskoncept, kan man tale om, at individer har en præference for frihed. Der er en række af eksperimentelle studier, der har taget den sidstnævnte tilgang og undersøger, om folk er villige til at betale for at få en større grad af frihed til selv at vælge, selv hvis dette sker på bekostning af deres materielle velfærd. Det har vist sig at være tilfældet (Bartling o.a., 2014; Bobadilla-Suarez o.a., 2017). Om end der også findes tilfælde, hvor forsøgsdeltagere foretrækker ikke selv at skulle bestemme over et givet udfald (Sunstein, 2014), f.eks. for at undgå skyld eller straf (Bartling og Fischbacher, 2012), lader det altså til at være tilfældet, at folk i mange

Figur 3: Frihed og velfærd i en adfærdsøkonomi



tilfælde sætter pris på autonomi og selvbestemmelse. Sunstein (2014) pointerer, at det kan være paternalistisk i sig selv at tvinge folk til at træffe et valg og at det til tider vil være ganske besværligt. Som eksempel kan her anvendes den amerikanske *Affordable Care Act*, hvor det er obligatorisk at vælge en sundhedsforsikring, og hvor man risikerer at modtage en bøde, hvis man vælger ikke at vælge. En oplagt mulighed er imidlertid, at man – frem for f.eks. at anvende automatisk tilmelding – giver folk mulighed for at vælge, samtidig med at man tilbyder en standard til dem, der ikke ønsker at vælge (Rebonato, 2012).

Dét, at folk sætter en værdi på frihed, er dog ikke nok til at vurdere om bestemte adfærdssinterventioner bør implementeres. En lidt naiv måde at undersøge det nærmere på er at anvende erklærede præferencer. Vi kan nærmere bestemt spørge folk, om de er for eller imod bestemte adfærdssinterventioner. Hvis folk går op i deres frihed, kan det formodes, at dette inkorporeres i deres overordnede vurdering af interventionerne. Grunden til, at dette er delvist naivt, er, som tidligere nævnt, at erklærede præferencer ikke er i stand til at tage højde for, at folk muligvis ikke er bevidste om deres adfærdsproblemer. Ikke desto mindre kan det give et fingerpeg om, om bestemte politiktiltag vurderes positivt eller ej, hvilket i sidste ende også spiller ind på sådanne tiltags politiske levedygtighed, jf. forløbet omkring indførelsen og afskaffelsen af den danske fedtskat.

Sunstein (forthcoming) byder på flere interessante konklusioner omkring, hvilke adfærdssinterventioner der har befolkningens opbakning. En af konklusionerne er, at der er større opbakning til nudges end mere traditionelle reguleringsformer. En årsag til dette kunne være, at interventioner er mere populære, hvis de formår at pille så lidt som muligt ved folks frihed, eller i hvert fald opfattes sådan. Sunstein finder også, at nudges, der

appellerer til folks bevidsthed og deliberative tankegange, har større opbakning end nudges, der appellerer til folks underbevidsthed og udnytter deres biases. Som eksempel viser folk større opbakning til madetiketter, der advarer om, at et produkt er usundt, end de gør til en intervention, der går ud på at ændre placeringen af varerne i supermarkeder, således at sunde varer får en mere fremtrædende position.

En ulempe ved ovenstående studier er, at de ikke skelner mellem de folk, som adfærdssinterventionen har til hensigt at hjælpe, og dem, som allerede handler optimalt. Fra et velfærds- og frihedsperspektiv er det den første gruppe, der er mest interessant, da det primært er denne gruppe, hvis velfærd og frihed er i spil. Lassen og Mahler (2017) forsøger at skelne mellem disse to gruppers holdninger. De viser, at de personer, som adfærdssinterventioner har til hensigt at hjælpe, systematisk er mere imod dem. Rygere er mere imod cigaretskatter, overvægtige er mere imod regulering af usund mad, folk der sparer lidt op, er mere imod forsøg på at få folk til at spare mere op osv. Dette er måske ikke særligt overraskende, da en delmængde af disse folk, som argumenteret for tidligere, rationelt underforbruger det bestemte gode. Hvis nogle har gode grunde til ikke at opspare, kan det forventes, at de er mere imod regulering, der tvinger dem til at opspare. Ovenstående resultat gælder dog også for folk, der er bevidste om deres adfærdsproblemer, såsom rygere, der erklærer, at de ønsker at stoppe med at ryge. Selv denne gruppe er mere imod adfærdssinterventioner end ikke-rygere. En årsag kunne være, at denne gruppe sætter pris på deres frihed og egen evne til at vælge.

Nudging og politisk økonomi

Slutteligt er det værd at knytte et par kommentarer til nudgingens politiske og administrative perspektiver. I midten af det tyvende århundrede var udgangspunktet for Buchanan og Tullocks (1962) kritik af den klassiske

velfærdsøkonomi, at man i velfærdsøkonomien på den ene side antog, at borgerne var rationelle, egeninteresserede væsener, mens embedsmænd og politikere antageligvis alene var motiveret af at gøre samfundet til et bedre sted. Public Choice-skolen, drevet frem af Buchanan og Tullock, antog omvendt, at alle agenter i den politiske proces – vælgere, embedsmænd, politikere – var egeninteresserede, antagelser som efterhånden er mainstream i store dele af moderne politisk økonomi.

Parallellen til nudging er lige til: Hvis bias og fejlbehæftet beslutningstagning er en menneskelig grundbetingelse, giver det så mening at tro, at politikere og embedsmænd kan sætte sig ud over sådanne egne skævheder i deres beslutningstagning? Er der forhold omkring den politiske proces, der gør, at bias og adfærdsøkonomiske problematikker konkurreres væk eller i sidste ende trumfes af rationel politisk deliberation? Den spæde litteratur på området er delt. Nogle studier viser, at den politiske proces forstærker, snarere end bremser, selvkontrolproblemer blandt vælgerne (f.eks. Bisin o.a., 2015), hvilket kan forstærke argumenter for second-best løsninger i form af lovkrav om balancerede budgetter. Andre, f.eks. Blumenthal (2012), accepterer, at eksperter, som tager beslutninger i professionaliserede institutionaliserede kontekster, også er underlagt kognitive bias, men påpeger, at den empiriske litteratur synes at pege på, at det gælder i væsentligt mindre grad for eksperter end for den almindelige borger. Blumenthal (2013) konkluderer derfor, at nok er eksperter også mennesker, og dermed ikke perfekte, men derfor kan de stadig bidrage til bedre politikbeslutninger end borgere overladt til sig selv – og deres egne valg. Det er under alle omstændigheder afgørende, at også dette metaniveau diskuteres med udgangspunkt i solide empiriske undersøgelser – hvis vi blot antager, at politikere og beslutningstagere kan afhjælpe den enkelte problemer med at vælge fornuftigt, er dette bare paternalisme med paternalisme på.

Noter

1. De tre forfattere er tilknyttet Økonomisk Institut og Center for Economic Behavior and Inequality (CEBI) på Københavns Universitet; David som professor, Daniel og Felix som ph.d.-studerende. David og Felix takker ERC Consolidator Grant nr. 313673 for finansiering. CEBIs aktiviteter er støttet af Danmarks Grundforskningsfond. Kontakt: ddl@econ.ku.dk.
2. Der er dog kreative løsninger, der gør det muligt at sammenligne folks velfærd. En ofte anvendt metode er blot at antage, at folk har ens præferencer. Denne metode anvendes som regel til beregning af optimal beskatning (Mirrlees, 1971). En anden metode er at kigge på såkaldt 'Kaldor-Hicks efficiens'. Her undersøger man, om et politisk tiltag kan gøre alle bedre stillet, hvis man samtidig laver en hypotetisk omfordeling mellem vindere og tabere. Denne tilgang ligger bag mange Cost-Benefit Analyser (Adler og Posner, 1999). Herudover kan man sammenligne folks velfærd ved at antage, at der er ét bestemt bundt, der giver alle individer den samme velfærd (Fleurbaey, 2009).

Referencer

- Adler, Matthew D. og Eric Posner (1999), »Rethinking cost-benefit analysis«, *The Yale Law Journal*, 109(2): 165-247.
- Bartling, Björn, Ernst Fehr og Holger Hertz (2014), »The Intrinsic Value of Decision Rights«, *Econometrica*, 82(6): 2005-39.
- Bartling, Björn og Urs Fischbacher (2012), »Shifting the Blame: On Delegation and Responsibility«, *Review of Economic Studies*, 79: 67-87.
- Bisin, Alberto, Alessandro Lizzeri og Leeat Yariv (2015), »Government policy with time inconsistent voters«, *American Economic Review*, 105(6): 1711-37.
- Blumenthal, Jeremy A. (2012), »Expert paternalism«, *Florida Law Review*, 64: 721-57.
- Blumenthal, Jeremy A. (2013), »A psychological defense of paternalism«, i: Christian Coons og Michael Weber, red., *Paternalism: Theory and practice*, New York: Cambridge University Press.
- Bobadilla-Suarez, Sebastian, Cass Sunstein og Tali Sharot (2017), »The intrinsic value of choice: The propensity to under-delegate in the face of potential gains and losses«, *Journal of Risk and Uncertainty*, 54(3): 187-202.
- Buchanan, James M. og Gordon Tullock (1962), *The*

- calculus of consent*, Ann Arbor: University of Michigan Press.
- Camerer, Colin, Linda Babcock og George Loewenstein (1997), »Labor Supply of New York City Cabdrivers: One Day at a Time«, *The Quarterly Journal of Economics*, 112(2): 407-41.
- Charness, Gary og Matthew Rabin (2002), »Understanding Social Preferences with Simple Tests«, *The Quarterly Journal of Economics*, 117(3): 817-69.
- Chetty, Raj o.a., (2014), »Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark«, *Quarterly Journal of Economics*, 129(3): 1141-1219.
- Chetty, Raj, Adam Looney og Kory Kroft (2009), »Salience and taxation: Theory and evidence«, *American Economic Review*, 99(4): 1145-77.
- Fleurbaey, Marc (2009), »Beyond GDP: The quest for a measure of social welfare«, *Journal of Economic Literature*, 47(4): 1029-75.
- Gneezy, Uri og John A. List (2006), »Putting Behavioral Economics to Work: Testing Gift Exchange in Labor Markets Using Field Experiments«, *Econometrica*, 74(5): 1365-84.
- Griffin, James (1986), *Well-being: Its meaning, measurement, and moral importance*, Oxford: Oxford University Press.
- Gruber, Jonathan H. og Sendhil Mullainathan (2005), »Do Cigarette Taxes Make Smokers Happier?«, *The BE Journal of Economic Analysis & Policy*, 5(1).
- Handel, Benjamin R. (2013), »Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts«, *American Economic Review*, 103(7): 2643-82.
- Handel, Benjamin R. og Jonathan T. Kolstad (2015), »Health Insurance for »Humans«: Information Frictions, Plan Choice, and Consumer Welfare«, *American Economic Review*, 105(8): 2449-2500.
- Kahneman, Daniel og Amos Tversky (1979), »Prospect theory: An analysis of decision under risk«, *Econometrica*: 263-91.
- Kahneman, Daniel, Peter Wakker og Rakesh Sarin (1997), »Back to Bentham? Explorations of Experienced Utility«, *The Quarterly Journal of Economics*: 375-406.
- Lassen, David D. og Daniel G. Mahler (2017), »Free to Choose or Free to Lose: Public Opinion over Government Paternalism«, *Working Paper*.
- Loewenstein, George og Drazen Prelec (1992), »Anomalies in intertemporal choice: Evidence and an interpretation«, *The Quarterly Journal of Economics*, 107(2): 573-97.
- Mas-Colell, Andreu, Michael D. Whinston og Jerry R. Green (1995), *Microeconomic theory*, New York: Oxford University Press.
- Mirrlees, James (1971), »An Exploration in the Theory of Optimum Income Taxation«, *The Review of Economic Studies*, 38(2): 175-208.
- O'Donoghue, Ted og Matthew Rabin (2003), »Studying optimal paternalism, illustrated by a model of sin taxes«, *American Economic Review*, 93(2): 186-191.
- Parfit, Derek (1984), *Reasons and persons*, Oxford: Oxford University Press.
- Poterba, James (2014), »Retirement security in an aging population«, *The American Economic Review*, 104(5): 1-30.
- Rebonato, Riccardo (2012), »A Critical Assessment of Libertarian Paternalism«, *Journal of Consumer Policy*, 37: 358-396.
- Rees-Jones, Alex. og Dmitry Taubinsky (2016), »Heuristic Perceptions of the Income Tax: Evidence and Implications for Debiasing«, *NBER Working Paper 22884*.
- Saint-Paul, Gilles (2011), *The tyranny of utility: Behavioral social science and the rise of paternalism*, Princeton: Princeton University Press.
- Simon, Herbert (1957), *Administrative Behavior: A study of decision-making processes in administrative organization*, Basingstoke: Macmillan.
- Stott, Henry og Ed Nottingham (2017), »Damage by Default: The Flaw in Pensions Auto Enrolment« i Alain Samson, red., *Behavioral Economics Guide 2017*. London.: Behavioral Science Solutions Ltd, pp. 32-40.
- Sunstein, Cass R. (forthcoming), »Do People Like Nudges«, *Administrative Law Review*.
- Sunstein, Cass R. (2014), »Choosing not to choose«, *Duke Law Journal*, 64(1): 1-52.
- Thaler, Richard. og Cass Sunstein (2008), *Nudge: Improving decisions about health, wealth, and happiness*, New Haven: Yale University Press.