

Om diskontering vid värdering av läkemedel

Johan Stennek*
Göteborgs universitet

2 maj 2021

*Jag är tacksam för synpunkter från Mats Bergman, Jonas Björnerstedt och Douglas Lundin.

”The truth is that the vast majority of real-world economics disputes are about easy questions—questions for which there is a clearly right answer . . . There are other, harder issues, like the question of which interest rate should be used in assessing infrastructure spending. But the easy questions provide plenty material to write about.“

Paul Krugman, 2020, *Arguing with Zombies: Economics, Politics, and the Fight for a Better Future*.

Innehåll

1	Sammanfattning	5
2	Varför diskontering?	8
3	Etiska utgångspunkter för diskontering	10
4	Samhällsekonomisk kostnads- och intäktsanalys	11
4.1	Nyttoteori	11
4.1.1	Några förenklande antaganden om separabilitet	12
4.1.2	Den rena tidspreferensen	14
4.1.3	Förväntad nytta och dödsrisker	15
4.2	Den samhällsekonomiska bedömningen	15
4.2.1	Kostnader	16
4.2.2	QALY	16
4.2.3	Betalningsvilja	17
4.2.4	Diskontering	18
5	Diskonteringsfaktorerna	18
5.1	Diskontering av kostnader	18
5.1.1	Offentliga sektorns (statens) låneränta	19
5.1.2	Marknadsimperfectioner	20
5.1.3	Korrelation mellan betalningar och ekonomisk aktivitet	20
5.1.4	Överensstämmelse med andra sektorer och “second best”	21
5.2	Diskontering av QALY	23
6	Diskontering med anledning av osäkerhet	24
A	QALY-beräkningar och förväntad nyttoteori	26
A.1	Qaly och betalningsvilja i en period	26
A.2	Diskontering	29
A.3	Upplåning och skattedistortioner	31
A.4	Osäkerhet (risk) rörande läkemedlens långsiktiga effekter	33
A.5	Osäkerhet om framtida materiell konsumtion och osäkerhet om hälsoeffekter	35
B	Den rena tidspreferensen som risk-aversion	36
C	Etiska utgångspunkter	37
C.1	Hälsa- och sjukvårdslagen	38
C.2	Intertemporala val ur ett livecykelperspektiv	40
C.3	Intertemporala val ur ett beslutsfattarperspektiv	44

C.4 Människovärdesprincipen enligt förarbetena	45
--	----

1 Sammanfattning

Denna PM demonstrerar att hälsoekonomiska värderingar av medicinska behandlingar endast överensstämmer med en samhällsekonomisk kostnads- och intäktsanalys om hälsovinster (QALY) och kostnader diskonteras. Faktum är att den diskonteringsfaktor som används har en avgörande betydelse för hur samhället värderar behandlingar som levererar hälsovinster under lång tid. Tyvärr är det, som citatet av Paul Krugman visar, inte helt enkelt att avgöra vad som är rätt diskonteringsfaktor.

Diskonteringsfaktorerna för QALY och kostnader ska bestämmas på olika sätt. Detta skiljer sig från gängse praxis, t.ex. på TLV.

Diskontering av hälsovinster (QALY) Diskonteringsfaktorn för QALY består av två komponenter. Den första komponenten är den representativa medborgarens "rena tidspreferens". Den andra komponenten beskriver hur värdet av ett år med perfekt hälsa utvecklas över tiden, till följd av ökad materiell levnadsstandard.¹

Den representativa medborgarens rena tidspreferens kan skattas empiriskt. Vi hittar dock inte några allmänt accepterade estimat. Vissa skattningar antyder en kraftig diskontering, men dessa kan vara behäftade med statistiska snedvridningar. I brist på mer tillförlitlig information förefaller en diskontering på ca en procent rimlig. Denna uppskattning bygger dock på ren introspektion. Detta skulle i så fall innebära en betydligt lägre diskontering av QALY än den TLV tillämpar idag.

Den rena tidspreferensen har en avgörande betydelse för hälsoekonomiska värderingar. Nedanstående tabell ger tre exempel på skillnader mellan en ren tidspreferens på tre procent, en procent (som förordas här) och noll procent.²

¹Hälsoekonomiska utvärderingar bygger på att man först räknar om hälsovinster i QALY (alltså i nytto-termer) och först därefter diskonterar hälsovinsterna. Detta förfarande innebär att diskonteringen blir något annorlunda än i samhällsekonomiska utvärderingar utanför hälsoområdet. På grund av att QALY-måttet i sig inbegriper aspekter så som dödsrisker och variationer i hälsa över livscykeln blir dessa faktorer inte en del av diskonteringen. Ett alternativt förfarande hade varit att först diskontera hälsoeffekter, för att därefter räkna om dem till QALY. Då hade diskonteringen även inkluderat dödsrisker och en justering av hälsoeffekterna med hänsyn till skillnader i marginalnyttan av hälsa i olika tidsperioder. Ett dylikt förfarande hade påmint mer om förfarandet i samhällsekonomiska utvärderingar utanför hälsoområdet.

²För att renodla betydelsen av den rena tidspreferensen bygger tabellen på en konstant BNP och att värdet av ett liv med perfekt hälsa därför inte förändras över tiden. Vidare bygger beräkningen av värdet av att rädda livet på en deterministisk livslängd på 83 år.

Värdet av att:		0%	1%	3%
1.	Förlänga livet med ett år för en femtioåring, jämfört med en nyfödd	100%	60%	20%
2.	Rädda livet på en femtioåring, jämfört med en nyfödd	40%	32%	16%
3.	Förlänga livet på en femtioåring med 33 år, jämfört med att förlänga livet på 33 femtioåringar med ett år	100%	55%	15%

Det tredje exemplet avser att illustrera att den rena tidspreferensen har särskilt stor betydelse för värdering av läkemedel som erbjuder bot i relation till värdering av läkemedel som förlänger livet. Dessa exempel visar att det vore önskvärt med ytterligare empiriska undersökningar för att fastställa storleken på människors rena preferens.

Den andra komponenten i diskontering av QALY är att värdet av ett år med perfekt hälsa ökar över tiden, till följd av ökad materiell levnadsstandard.³ Vi uppskattar denna ökning mycket grovt till ca en procent. Denna uppskattning baseras på ett antagande om en BNP per capita på ca 500 tusen kronor, en tillväxt på ca 2% om året, samt att betalningsviljan för ett år med perfekt hälsa är ca 1 miljon kronor.

Om vår gissning att den rena tidspreferensen föranleder en diskontering på en procent och att det ökade värdet av ett liv i perfekt hälsa föranleder en uppskrivning med en procent, blir nettoeffekten att QALY inte ska diskonteras. Denna slutsats innehåller en ironi. Den ursprungliga QALY-modellen baseras på en nyttofunktion som endast beror på personens hälsa.⁴ Den bortser ifrån den rena tidspreferensen och även ifrån materiell konsumtion. När vi lägger till dessa faktorer visar det sig att de nettar ut varandra och att den enklare ursprungliga QALY-modellen därför kan vara en god approximation.

Diskontering innehåller etiska ställningstaganden. Att ta hänsyn till medborgarnas egna preferenser, såväl deras "rena tidspreferens" som att deras betalningsvilja för hälsa torde öka till följd av ekonomisk tillväxt, innebär att QALY skulle kunna komma att värderas olika beroende på kronologisk ålder. Detta illustreras särskilt av första raden i tabellen ovan. Ett sådant förfarande är enligt vår uppfattning förenligt med hälso- och sjukvårdens människovärdesprincip. Förfarandet är besläktat med ett pensionssystem där yngre betalar äldres pensioner. Det innebär vinster för alla generationer, sett över livscykeln. Enligt en etablerad tolkning av människovärdesprincipen får dock QALY inte värderas olika med hänsyn till kronologisk ålder. Det skulle i så fall innebära att QALY inte bör diskonteras alls. I praktiken spelar detta dock inte någon roll eftersom även vi kommer fram till att QALY inte ska diskonteras, om än av andra skäl.

³Inom traditionell diskontering tar man hänsyn till att marginalnyttan av konsumtion kan förmodas minska över tiden till följd av tillväxten. När man diskonterar QALY är det tvärt om. Tillväxt innebär ökad betalningsvilja för hälsa och att framtida hälsovinster därför ska räknas upp.

⁴Se t.ex. Milton C. Weinstein, George Torrance, Alistair McGuire: QALYs: The Basics. Volume 12, Supplement 1, 2009, *Value in Health*.

Diskontering av kostnader Utgångspunkt för diskontering av kostnaderna bör vara statens upplåningsränta.⁵ Denna slutsats bygger på att samhället (staten) fritt kan fördela kostnader mellan olika år genom lån. Låneräntan kan dock behöva justeras av olika anledningar.

Möjligen behöver upplåningsräntan justeras för att ta hänsyn till imperfektioner på kapitalmarknaden. Dessa är att staten eventuellt besitter monopsonikraft på lånemarknaden samt att statliga investeringar eventuellt undantränger privata investeringar med ett socialt värde som är högre än investerarens privata värde.

Vi menar däremot att justering för att ta hänsyn till korrelationer mellan läkemedelsbetalningar och den ekonomiska aktiviteten sannolikt inte är påkallad. Läkemedel är skattefinansierade och riskerna är därmed diversifierade. Vidare torde en korrelation mellan läkemedelsbetalningar och aktivitetsnivån i ekonomin saknas.

Sammantaget innebär detta att diskonteringen av kostnader för läkemedel sannolikt bör vara lägre än den diskontering som TLV tillämpar idag.

Diskonteringen skulle även vara lägre än diskontering av t.ex. statliga infrastrukturinvesteringar (t.ex. höghastighetståg). Diskontering av dylika investeringar handlar i främsta rummet om avlägsna intäkter, som på goda grunder kan förmodas samvariera med konjunkturläget.⁶

Frågan är om skillnader i diskonteringsfaktorer mellan olika verksamheter i sig leder till några problem, och om TLV av denna anledning bör använda samma diskontering som andra offentliga aktörer. Vi menar att den diskonteringsfaktor som användas för värdering av läkemedel inte ska justeras av denna anledning. Diskonteringsfaktorn ska användas för prioritering av projekt över tiden inom en verksamhet (t.ex. prioritering mellan ATMP och vanliga läkemedel). Behöver man av någon anledning justera ned den totala läkemedelskostnaden (t.ex. på grund av att man använder en lägre kalkylränta än andra) är det bättre att justera ned betalningsviljan för QALY. Denna ska användas för att prioritera mellan olika verksamheter, så som läkemedel och väginvesteringar.⁷

⁵Denna utgör ett mått på den rena tidspreferensen och utvecklingen av marginalnyttan av materiell konsumtion.

⁶En studie av hur diskonteringsfaktorer sätts i ett antal länder i norra Europa antyder att svenska diskonteringsfaktorer bestäms genom förhandling. Den diskontering som används inom transportområdet är fastställd i en förhandling mellan trafikverket, som förespråkar en låg ränta vilket talar för fler projekt, och finansdepartementet, som förespråkar en hög ränta vilket kan förmodas leda till lägre offentliga utgifter. Se Mo U ter, n . (2018). A critical assessment of discounting policies for transport cost-benefit analysis in five European practices. *European Journal of Transport and Infrastructure Research*, 18(4), 389–412.

⁷Om frågan rör prioriteringar inom sjukvården är det bättre att åsätta olika betalningsvilja för QALY inom olika delar av vården. QALY är skapade för att fördela resurser mellan sjukvård och andra sektorer och även för att fördela resurser mellan olika verksamheter inom sjukvården. Utgångsläget är då att alla inom sjukvården bör ha samma QALY. Om delar av sjukvården använder felaktiga diskonteringsfaktorer är det dock bättre att justera betalningsviljan för QALY inom läkemedelsvärderingar än att justera diskonteringsfaktorn.

Osäkerhet För nya läkemedel, vars långtida hälsovinster fortfarande är osäkra, beräknas *förväntade* QALY i varje tidpunkt, som därefter diskonteras på samma sätt som när effekterna är säkra.⁸

När ett läkemedels långsiktiga effekter är osäkra kan det finnas anledning att använda en betalningsmodell där betalningarna till läkemedelsbolagen betingas på utfallet, om det är praktiskt möjligt att mäta detta. Oavsett om betalningarna betingas på utfallet bör inte någon risk-justering ske, försvärd inte behandlingen ges till en stor andel av befolkningen.⁹ Orsaken är för det första att betalningarna är skattefinansierade och att risken därför är väl diversifierad och, för det andra, att betalningarna inte kan förmodas vara korrelerade med aktivitetsnivån i ekonomin.

2 Varför diskontering?

En viktig skillnad mellan ATMP och många andra läkemedel är behandlingseffektens längd. En ATMP-behandling förmodas leverera hälsovinster för patienten långt in i framtiden och möjligen erbjuda bot. Effekten av vanliga läkemedel upphör ofta efter en tid och behandlingen behöver upprepas med jämna mellanrum. Frågan är då om vi idag ska använda sjukvårdens budget till behandlingar som ger oss omedelbara hälsovinster eller, alternativt, för behandlingar som framförallt ger upphov till hälsovinster i framtiden. För att kunna välja måste vi värdera hälsovinster vid olika tidpunkter mot varandra. Med en hälsovinst menar vi t.ex. att slippa en viss specifik smärta under ett års tid. Det är här den så kallade diskonteringsfaktorn kommer in. Den definieras som det värde som vi idag åsätter att slippa smärtan först under nästa år i relation till det värde som vi åsätter att slippa samma smärta redan i år. Vi kommer att visa hur denna diskonteringsfaktor kan bestämmas. Och den grundläggande frågan är varför inte hälsovinster ska anses lika värdefulla, oavsett när de uppstår? Vad finns det för anledning att värdera framtida hälsovinster lägre, eller kanske högre, än de hälsovinster som vi kan åstadkomma idag?

Vår huvudfråga avser värdering av hälsovinster för en och samma person (eller åtminstone personer inom samma generation) men som uppträder vid olika tidpunkter. Men diskontering används även för att jämföra hälsovinster som tillfaller framtida generationer med de hälsovinster som tillfaller de som lever nu. Detta är dock fråga om två olika typer av värderingar. Diskontering mellan olika generationer består huvudsakligen i att värdera olika människoliv mot varandra. Vad som är en lämplig diskonteringsfaktor i det ena fallet behöver inte vara lämplig i det andra fallet. Även om diskontering mellan

⁸QALY-beräkningen tar i sig hänsyn till att nyttan av ökad livskvalitet kan öka i avtagande takt (konkavitet). Av detta följer att ingen riskjustering ska göras.

⁹Om betalningarna betingas på utfallet *ökar* det offentliga risktagandet. Hälso risken med att ett läkemedel kanske inte fungerar bär patienterna, alltså de som får läkemedlet. Den finansiella risk som uppkommer om betalningen för läkemedlet betingas på utfallet bär skattebetalarna. Eftersom patienter och skattebetalare i stort sätt är olika personer (om inte läkemedlet ges till en stor andel av befolkningen) innebär betingningen inte att lägre betalningar kompenserar för ett dåligt hälsoutfall.

generationer är en bisak för vårt syfte, kan vi inte undgå att beröra frågan. Orsaken är som följer. Människor åsätter ett lägre (eller ibland högre) värde för de hälsovinster som kommer längre fram i deras liv. Om samhället accepterar dessa värderingar, får det följder för hur samhället måste värdera människor med olika ålder som lever i samma tidpunkt. Detta reser etiska spörsmål som vi behandlar i ett särskilt avsnitt.

En andra skillnad mellan ATMP-läkemedel och andra läkemedel är när betalningen sker. Betalning för behandling med vanliga läkemedel som upprepas kontinuerligt måste normalt erläggas kontinuerligt vid leverans. Betalning för ett ATMP-läkemedel kan göras i form av en engångsbetalning i samband med behandlingen. Men man kan även komma överens om att sprida ut betalningen över tiden. Därmed uppkommer frågan om hur vi ska värdera en stor betalning idag mot en ström av mindre betalningar i framtiden. Återigen är svaret diskontering. Dock behöver diskonteringsfaktorn för hälsovinster och för betalningar inte vara densamma.

Betalning för läkemedel innebär att patienten, i sin egenskap av skattebetalare, måste avstå från annan materiell konsumtion. En betalning i år innebär att patienten till exempel kan äta mindre potatis i år. (Även om ett konkret exempel som potatis kan förefalla lite löjligt tror vi att resonemangen blir tydligare.) En betalning nästa år innebär att patienten kan äta mindre potatis nästa år. Det relativa värdet av potatis nästa år i förhållande till idag behöver inte vara detsamma som det relativa värdet av en hälsovinst nästa år i förhållande till samma hälsovinst idag. Om vi fortsätter att bli rikare kommer vi att konsumera mer potatis i morgon än idag. Eftersom vi kommer att vara "mättare" i morgon kommer förlusten av en potatis då att kännas mindre. Värdet av en potatis i morgon är således ganska lågt i förhållande till värdet av en potatis idag. Med hälsa är det sannolikt tvärt om eftersom hälsan sjunker med åldern. Att förbättra sin kondition (maximal syreupptagningsförmåga per kilo) med en enhet från en låg nivå kan vara mer värdefullt än att förbättra den med en enhet från en redan hög nivå. Således kan samma hälsovinst vara mer värdefull vid äldre år.

En tredje skillnad mellan ATMP och redan existerande läkemedel är att ATMP är nya. De är därför behäftade med betydligt större osäkerheter rörande deras effekter. Förhoppningen är att ATMP ska ge hälsovinster under lång tid och kanske bot. Det kan man förstås inte veta förrän lång tid efter deras lansering. Och människor ogillar osäkerhet. Annorlunda uttryckt värderar man två alternativa behandlingar med samma förväntade hälsoeffekter olika beroende på hur osäkra effekterna är. De flesta skulle föredra en behandling som ökar deras rörlighet med tio procent med säkerhet framför en som ökar rörligheten med tjugo procent eller inte alls med lika sannolikhet. Framtida hälsovinster måste således "diskonteras" mer om de är osäkra. Dock är denna typ av diskontering redan inkorporerad som en del av själva QALY-beräkningen. QALY är ju inte något annat än nyttan av hälsoeffekten. Genom att diskontera förväntade QALY (alltså förväntad nytta) behövs ingen risk-korrigerings.

3 Etiska utgångspunkter för diskontering

Vi förordar och använder en sedvanlig samhällsekonomisk kostnads-intäktsanalys för att värdera ATMP-läkemedel. Denna ansats bygger på förväntad nyttoteori och utilitarism. Vi menar att ansatsen överensstämmer med och preciserar den etiska plattform som ska styra sjuk- och hälsovården i Sverige.¹⁰

Den etiska plattformens första princip, människovärdesprincipen, anses vara överordnad de övriga. Den tolkas ofta utslutande “negativt”, att den stadgar vilka omständigheter som *inte* får läggas till grund för prioriteringar. Dock består människovärdesprincipen av två delar. Den första delen är att alla personer värderas lika. Det är denna likavärdesprincip som föreskriver att prioriteringar *inte* får baseras på en mängd olika faktorer såsom inkomst och socio-ekonomisk status. Och, applicerad på diskontering, menar vi att likavärdesprincipen innebär att personer som tillhör olika generationer ska behandlas lika.¹¹

Den andra delen av människovärdesprincipen stadgar att samhällets prioriteringar bör överensstämma med medborgarnas egna. Denna så kallade integritetsprincip är “positiv”. Om människor måste välja och föredrar att bli av med en viss smärta framför att öka sin rörlighet, så ska en behandling mot smärtan värderas högre än en som ökar rörligheten, allt annat lika. Integritetsprincipen innebär därmed att empiriska studier av människors egna värderingar bör vara det främsta instrumentet för prioriteringar inom vården. Applicerad på diskontering, menar vi följaktligen att integritetsprincipen föreskriver att samhället bör diskontera framtida hälsovinster i enlighet med medborgarens egna tidspreferenser. Och empiriska studier visar bl.a. att människor har så kallade “rena tidspreferenser”. Det innebär att de, allt annat lika, värderar en viss nytthet allt lägre ju längre fram i tiden nyttheten förväntas erhållas. Om en person räknar med att vara vid samma hälsa under innevarande år och nästkommande år, och om personen även räknar med oförändrad inkomst och andra levnadsvillkor övrigt, och om personen får välja att förlägga en (odelbar) hälsovinst till innevarande eller nästkommande år, så föredrar personen det tidigare alternativet.

Rörande diskontering kan människovärdesprincipen således preciseras på följande sätt:

1. Likavärdesprincipen: Hälsovinster som tillfaller en (säg) 70-åring i morgon ska värderas på samma sätt som om dessa hälsovinster hade tillfallit en 70-åring (med samma hälsotillstånd m.m.) idag.
2. Integritetsprincipen: Hälsovinster som tillfaller en och samma person men vid olika åldrar ska värderas utifrån personens egna tidspreferenser.

¹⁰Analysen presenteras mer utförligt i Appendix C.

¹¹Det förefaller även finnas en enighet bland ekonomer att det saknas en etisk grund för diskontering av framtida generationers nytta. Notera dock att det är fullt möjligt att diskontera framtida generationers konsumtion på grundval av förväntad tillväxt och risken för mänsklighetens utdöende (vilket varit en del av debatten rörande diskontering i klimatdebatten). Se diskussionen i Gollier, sid. 31.

Problemet är att dessa båda principer implicerar följande slutsats:

- Om människor själva värderar hälsovinster lägre (eller högre) ju längre fram i livet de kommer, bör samhället värdera hälsovinster som tillfaller äldre personer idag lägre (eller högre) än motsvarande hälsovinster som tillfaller yngre personer idag.

Denna slutsats är kontroversiell eftersom den kan synas vara diskriminerande. Vi menar att detta dock endast är skenbart diskriminerande eftersom de som är gamla idag var unga i går och de som är unga idag blir gamla i morgon. Över livscykeln behandlas alla lika.¹²

Detta förfarande kan t.o.m. möjligen ses som ett bidragande skäl till varför offentligt finansierad sjukvård kan vara mer fördelaktig för medborgarna än en marknadslösning. Varje generation får en fördel som ung av de som är gamla då och återgäldar senare denna fördel som gamla till de som är unga då. Det är inte uppenbart att en dylik allokering kan uppnås via enskilda avtal på marknaden, eftersom inte alla lever på samma gång.¹³

Våra slutsatser strider dock mot den tolkning av människovärdesprincipen som redovisas i förarbetena (prop. 1996/97:60). Enligt förarbetena ska en viss hälsovinst i en viss tidpunkt värderas på samma sätt oberoende av om den tillfaller en äldre eller en yngre person. Om denna tolkning kombineras med ett krav på att hälsovinster som uppträder i olika tidpunkter ska värderas lika (vilket är en förutsättning för att olika generationer ska värderas lika), följer det att man vid diskontering av hälsovinster och kostnader för dessa inte ska beakta den rena tidspreferensen.

4 Samhällsekonomisk kostnads- och intäktsanalys

4.1 Nyttoteori

Utgångspunkten för den samhällsekonomiska analysen är hur en medicinsk behandling och dess betalning påverkar en representativ medborgare, såväl i egenskap av patient som skattebetalare. Det är den representative medborgarens egna preferenser som ligger till grund för bedömningen. Dessa preferenser beskriver vi med en nyttofunktion. Säg att medborgaren lever i två tidsperioder och att denne endast bryr sig om två omständigheter, sin egen hälsa och sin egen materiella konsumtion. Då skriver vi

$$\tilde{U}(H_0, C_0, H_1, C_1),$$

¹²Det finns flera andra exempel på att politiska beslut måste ses i ett livcykelperspektiv för att inte framstå som åldersdiskriminerande. Ett mycket närliggande exempel är pensionssystemet som innebär att de som är unga idag betalar pensionerna för de som är gamla idag, mot ett löfte om att få motsvarande behandling i äldre år. Detta kallas pay-as-you-go och bygger på liknande tankar som vi presenterar här. Ett annat exempel är den värnplikt som drabbar varje generation som ung och som är till fördel även för den generation som då är gammal.

¹³En person i generation t får prioriterad sjukvård som bekostas av en person i generation $t - 1$, men återgäldar till en person i generation $t + 1$, som kanske inte ens är född vid det första tillfället.

där H_t är någon aspekt på hälsa t.ex. kondition (maximal syreupptagningsförmåga) och C_t anger personens materiella konsumtion i de två tidpunkterna. Enklast är att tänka att konsumtion av andra varor mäts i kronor.

Vi föreställer oss att den representativa medborgaren föredrar mer hälsa och mer konsumtion i varje tidpunkt. Annorlunda uttryckt är marginalnyttorna av både hälsa och konsumtion positiv. Vi föreställer oss även att varje ytterligare tillskott av hälsa eller konsumtion ger ett allt mindre nyttotillskott. Annorlunda uttryckt är marginalnyttorna avtagande.

Nyttofunktionen kan tolkas på olika sätt. Enligt den hedonistiska tolkning beskriver nyttofunktionen det välbefinnande som en hungrig person känner när den äter. Enligt en annan tolkning beskriver nyttofunktionen även mer abstrakta preferenser.¹⁴ Ett exempel skulle kunna vara att avstå ifrån att dricka alkohol av religiösa skäl, även om det skulle ge personen samma typ av välbefinnande som en ateist.

4.1.1 Några förenklande antaganden om separabilitet

För att beskriva människors val och välbefinnande över tiden kommer vi att göra några förenklande antaganden. Dessa antaganden är standard i tillämpad ekonomi. De är dessutom standard i hälsoekonomiska utredningar som baseras på QALY.

Separabilitet över tiden Det första antagandet är att nyttofunktion är separabel över tiden. Den kan då skrivas

$$\tilde{U}(H_0, C_0, H_1, C_1) = \tilde{U}^0(H_0, C_0) + \tilde{U}^1(H_1, C_1).$$

Det betyder att vi bortser ifrån att vårt välbefinnande av morgondagens konsumtion C_1 inte påverkas av dagens konsumtion C_0 . Så behöver inte vara fallet. Ett motexempel är vanebildande konsumtion.

Det andra förenklande antagandet är att sambandet mellan å ena sidan hälsa och konsumtion i en viss tidpunkt och å andra sidan nytta ser ut på samma sätt vid båda tidpunkterna. Mer precist föreställer vi oss att nyttofunktionen kan skrivas

$$\tilde{U} = U(H_0, C_0) + \delta \cdot U(H_1, C_1).$$

där δ uttrycker personens så kallade rena tidspreferens. Empiriska studier visar att den rena tidspreferensen är lägre än ett, $\delta < 1$. Frågan är nu hur detta uttryck ska tolkas.

1. Med en rent hedonistisk tolkning skulle den rena tidspreferensen innebära att välbefinnandet av att äta middag avtar med stigande ålder. Denna tolkning förefaller oss vara mer falsk än sann.

¹⁴Harsanyi

2. En annan vanlig tolkning av den rena tidspreferensen är att den uttrycker människors otålighet. Även om välbefinnandet av att äta middag är detsamma i båda tidpunkterna, föredrar man ett omedelbart välbefinnande framför ett senarelaggt.
3. En tredje tolkning, som vi kommer att utveckla nedan, är att den rena tidspreferensen uttrycker personens aversion mot risker i livslängd.

Vi kommer här i huvudsak uttrycka oss i termer av den tredje tolkningen. Ett möjlig invändning mot denna tolkning (och även den andra) är att den förutsätter ett slags extremt *ex ante* perspektiv, att personen som nyfödd planerar hela sitt liv och värderar t.ex. olika hälsotillstånd över livscykeln. Det är förstås en abstraktion. Men människor gör intertemporala val, t.ex. när de sparar. Alla dessa intertemporala val görs per definition utifrån ett *ex ante* perspektiv (om än mer begränsat) och de vägleds av deras tidspreferenser.¹⁵

Vi förutsätter att samhällets värderingar överensstämmer med den representative medborgarens värderingar. Annorlunda uttryckt förutsätter vi att en offentlig beslutsfattare i egenskap av agent för den representative patienten fattar de beslut som patienten själv skulle vilja fatta om denne t.ex. hade haft tillgång till all tillgänglig information.

Att många val så som sparande är intertemporala innebär att den rena tidspreferensen även kan skattas empiriskt genom att studera dylika val.¹⁶ Det bör således vara fullt möjligt att även i praktiken basera hälsoekonomiska bedömningar på medborgarnas tidspreferenser.¹⁷

Separabilitet mellan hälsa och konsumtion Syftet med vår analys är att den ska sättas in i ett hälsoekonomiskt sammanhang där hälsovinster värderas med hjälp av QALY. Såvitt vi förstår räknar med samma QALY för samma hälsovinst, oavsett om häl-

¹⁵Man kan även invända att preferenserna inte behöver vara konstanta. Vad är det som säger att de inte ändras? En starkt tidspreferens kanske gör att man röker som ung, och ångrar det som gammal. Denna invändning skakar grunden för nyttoteorin och idén om rationella val. Dessvärre är dock för närvarande nära nog all ekonomisk analys begränsad av dessa antagande.

¹⁶Experimentella studier av intertemporala val visar att försökspersoner föredrar att erhålla en viss belöning omedelbart framför att erhålla samma belöning (säg) ett år senare. Däremot är samma personer mindre benägna att idag göra en åtskillnad mellan en belöning om tio och en belöning om elva år. Det förefaller alltså som att diskonteringen mellan två på varandra följande perioder i närtid är högre än diskonteringen mellan samma två på varandra följande perioder i framtiden. Detta fenomen benämns hyperbolisk diskontering. Evidensen visar att hyperbolisk diskontering endast påverkar val som ligger nära i tiden, inom ett år (Frederick et al, i *Advances in Behavioral Economics*, red Camerer et al.). Det betyder att den hyperboliska diskonteringen inte kommer att ha någon kvantitativ betydelse för värdering av läkemedel. Vi bortser därför ifrån denna komplikation. Man kan även ställa frågan om en offentlig beslutsfattare, i egenskap av agent för den representative medborgaren, borde anamma detta förfarande oavsett tidshorisont. Frågan måste ställas eftersom hyperbolisk diskontering leder till tidsinkonsistenta val och möjligen är att betrakta som ett uttryck för en begränsad rationalitet. Vi menar att en offentlig beslutsfattare därför inte bör anamma förfarandet. Detta förfaller vara rimligt på samma sätt som att en offentlig beslutsfattare i egenskap av agent för den representative patienten kan göra bättre val å patientens vägnar till följd av bättre information om t.ex. läkemedels effekter.

¹⁷Att QALY baseras på separabilitet över tiden framgår av att QALY, i sin mest grundläggande form, beräknas genom att summera nyttan i olika hälsotillstånd, multiplicerade med antal år individen befinner sig i dessa tillstånd.

sovinsten uppstår idag eller i framtiden när den materiella välfärden kan förväntas vara högre. För att detta ska vara ett korrekt förfarande krävs det att nyttofunktionen i en viss tidpunkt är (multiplikativt) separerbar i hälsa och materiell konsumtion. Det betyder att den ges av

$$U(H_t, C_t) = u(H_t) \cdot v(C_t).$$

Vi kommer att beskriva detta närmare i nästa avsnitt. Vi kommer dock fortsätta att skriva $U(H_t, C_t)$ och redovisa explicit när vi använder separabiliteten mellan hälsa och materiell konsumtion.

4.1.2 Den rena tidspreferensen

Den rena tidspreferensen (δ) förknippas ofta med otålighet. Vissa menar att människor inte *borde* vara otåliga, d.v.s. att människors rena tidspreferens *borde* vara lika med ett. Ramsey menade t.ex. att diskontering orsakas av människors bristande förmåga att leva sig in i deras framtida jag. Andra har beskrivit den rena tidspreferensen som känslornas vinst över förnuftet.¹⁸

Men den rena tidspreferensen kan även tolkas på ett annorlunda och kanske mer fundamentalt sätt. Den beskriver nämligen den representative medborgarens aversion mot risker avseende livslängd (se Appendix B). Ett hypotetiskt exempel belyser detta påstå-
ende. Säg att en person med en allvarlig sjukdom utan behandling kommer att avlida om exakt ett år. Säg att det finns en behandling som antingen leder till omedelbar död eller till X extra års liv med lika sannolikhet. Vi tänker oss för enkelhets skull att personen har god hälsa fram till dödsögonblicket oavsett när det inträffar. Mer precist tänker vi oss att livskvaliteten är densamma. Hur många extra år X skulle behandlingen behöva ge för att patienten skulle vilja ta del av behandlingen? Det finns studier som antyder att människor inte nöjer sig med $X = 2$ extra år. Om $X > 2$ så föredrar man ett års liv med säkerhet framför den riskfyllda behandlingen, trots att den förväntade livslängden då skulle bli längre. Säg att personen precis ska till att fylla 60 år. Vi betecknar värdet av det kommande året med u_{60} och värdet av det därefter följande året med u_{61} .¹⁹ Då är personens nytta utan behandling just u_{60} . Den förväntade nyttan med behandlingen blir $\frac{1}{2} \cdot (u_{60} + u_{61})$, eftersom personen överlever behandlingen med 50% chans och då får två extra år. Att behandlingen har lägre värde än att avstå betyder att $\frac{1}{2} \cdot (u_{60} + u_{61}) < u_{60}$. Det betyder i sin tur att $u_{61} < u_{60}$. Att ogilla risk rörande livslängd är med andra ord liktydigt med att varje år är mer värdefullt än varje nästkommande år. Det är kvoten $\frac{u_{61}}{u_{60}} < 1$ som benämns ren tidspreferens. Vi betecknar den δ och kommer för enkelhets skull förutsätta att den är konstant över livet.²⁰

¹⁸Se diskussionen i Gollier, sid. 31.

¹⁹Med den notation vi hade ovan $u_{60} = \tilde{U}^0(H, C)$ och $u_{61} = \tilde{U}^1(H, C)$.

²⁰Den rena tidspreferensen påverkar värderingen av framtida hälsovinst. Ett enkelt exempel belyser detta påstående. Den nytta personen får under ett år beror bland annat på hälsan under året. Om detta

Vår bedömning är att den rena tidspreferensen är av stor kvantitativ betydelse för värdering av ATMP. Det finns t.o.m. empiriska studier som antyder att den rena tidspreferensen kan vara mycket kraftig. I en studie²¹ av patienter med testikelcancer fann man att medianen bland respondenterna var indifferent mellan 4 år med god hälsa med säkerhet och en behandling som orsakar omedelbar död med 50% sannolikhet och ett liv med god hälsa under 10 år med 50% sannolikhet. Detta motsvarar årlig diskonteringsfaktor $\delta = 0.92$. Det vill säga att varje år är 8% mindre värt än varje föregående år. Vår gissning är dock att denna skattning är kraftigt överdriven. Vi kommer därför att använda oss av en betydligt lägre ren tidspreferens för våra illustrativa beräkningar. Vi kommer närmare bestämt välja 1% eller annorlunda uttryckt att $\delta = 0,99$. Även denna lägre diskontering får en stor inverkan på värdering av läkemedel. Det innebär, allt annat lika, att en hälsovinst om 15 år bara är värd 85% av samma hälsovinst idag.

4.1.3 Förväntad nytta och dödsrisker

Vi behöver även värdera framtida hälsovinster när det finns en risk för att den representativa patienten hinner avlida innan dessa hälsovinster uppstår. Denna risk kan vara av vilket slag som helst, alltså både risker som är kopplade till redan uppkomna sjukdomar eller t.ex. risken för en trafikolycka. För att ta hänsyn till risken att dö låter vi p_t beteckna sannolikheten att den representativa patienten lever i period t . Förväntad nytta ges då av

$$E\tilde{U} = p_0 \cdot U(H_0, C_0) + \delta \cdot p_1 \cdot U(H_1, C_1).$$

Vi normaliserar patientens nytta till att vara lika med noll om denne är död.

4.2 Den samhällsekonomiska bedömningen

För att avgöra om en medicinsk behandling är samhällsekonomiskt motiverad jämför vi kostnaderna för behandlingen med intäkterna. Vi börjar med att studera en behandling som bara har medicinska effekter under en period och som betalas i samma period.

samband är konstant över tiden kan vi beskriva det med en nyttofunktion $u_t = U(H_t)$ där H_t betecknar hälsotillståndet under året. Således får en person som har samma hälsa under två på varandra följande år en total nytta som ges av $U(H) + \delta \cdot U(H)$. Om personen får en liten hälsovinst under det första året är värdet marginalnyttan av hälsa, $U'(H)$. Om personen får en liten hälsovinst under det andra året så är värdet idag endast $\delta \cdot U'(H)$. Värdet av en hälsovinst nästa år i förhållande till värdet av samma hälsovinst redan i år ges med andra ord av den rena tidspreferensen, $\frac{\delta \cdot U'(H)}{U'(H)} = \delta < 1$.

²¹Stiggelbout, A. M., Kiebert, G. M., Kievit, J., Leer, J. W. H., Stoter, G. and de Haes, J. C. J. M. (1994). 'Utility assessment in cancer patients: adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores', Medical Decision Making, vol. 14, pp. 82–90. (Vår rapportering genom: Han Bleichrodt and Jose Luis Pinto: THE VALIDITY OF QALYS UNDER NON-EXPECTED UTILITY. The Economic Journal, 115 (April), 533–550.)

4.2.1 Kostnader

Kostnaderna utgörs såväl av det pris läkemedelsföretaget begär som den resursåtgång som krävs i vården för att administrera behandlingen.²² Dessa kostnader betalas av medborgarna om än via skatten. De leder därmed till att konsumtion av varor och tjänster, d.v.s. C_t i nyttofunktionen ovan, minskar. Om vi i beräkningen antar att det är samma person som får behandlingen och som ska betala kostnaden måste vi ta hänsyn till risken att personen dör. Det gör vi genom att låta personen betala $dC_t = -\frac{\text{Kostnad}_t}{p_t}$ i det fall personen överlever. Då blir den förväntade betalningen lika med kostnaden.²³

4.2.2 QALY

Intäkterna utgörs av hälsovinster i form av ökad överlevnad eller ökad livskvalitet, t.ex. ökad kondition eller minskad smärta. Hälsovinster räknas om till kvalitetsjusterade levnadsår, QALY, innan de diskonteras eller jämförs med kostnaderna. QALY-måttet har sin grund i nytte teorin. Det baseras dock på ett antal förenklande antaganden om nyttefunktionens utseende, bland annat att nyttefunktion är separabel över tiden. Den ursprungliga QALY-modellen förutsätter även att människor är risk-neutrala avseende antal levnadsår samt att värdet av hälsa inte påverkas av materiell levnadsstandard. Dessa två antaganden implicerar att ingen diskontering ska ske.²⁴ Idag använder dock många myndigheter inom hälsoområdet diskontering.²⁵

QALY-måttet baseras vidare på tanken att nyttan av att vara död är noll och att det även finns en högsta möjlig nytta, nämligen nyttan vid perfekt hälsa, vilket åsätts värdet ett. Varje hälsotillstånd kan därefter ges ett värde som anger nyttan i relation till nyttan vid perfekt hälsa. En person som har hälsan H_t och materiell konsumtion C_t i tidpunkt t åtnjuter nytta $U(H_t, C_t)$. Hade denne person istället varit vid perfekt hälsa, H^* , hade nyttan varit $U(H^*, C_t)$. Det innebär att personens nytta, uttryckt i QALY, ges av $\frac{U(H_t, C_t)}{U(H^*, C_t)}$. Om vi använder antagandet om (multiplikativ) separabilitet mellan hälsa och materiell konsumtion blir uttrycket $\frac{U(H_t, C_t)}{U(H^*, C_t)} = \frac{u(H_t)}{u(H^*)}$. En anledning till att QALY-måttet är attraktivt är sannolikt att det är enklare att fråga människor hur de värderar sitt liv i ett visst hälsotillstånd H_t , om man sätter detta i relation till ett liv i perfekt hälsa.

QALY-måttet används för att mäta hälsovinster, t.ex. i form av ökad chans att överleva

²²En del av det pris som betalas till läkemedelsföretagen utgörs av ett vinstpåslag och skulle kunna betraktas som en transferering snarare än en samhällsekonomisk kostnad. Vi anammar dock ett svenskt perspektiv och eftersom flertalet företag är utländska är betalningen att betrakta som en kostnad.

²³Alternativt kan vi räkna med att andra personer betalar kostnaden. Detta ger samma resultat och redovisas i Appendix.

²⁴Se t.ex. Milton C. Weinstein, George Torrance, Alistair McGuire: QALYs: The Basics. Volume 12, Supplement 1, 2009, *Value in Health*.

²⁵En översikt av nationella riktlinjer för diskontering i hälsoekonomiska utvärderingar i 24 länder visar att man använder diskonteringsfaktorer mellan 0% och 5% och att de vanligaste värdena är 5% och 3%. Attema, A.E., Brouwer, W.B.F., Claxton, K.: Discounting in economic evaluations. *Pharmacoeconomics* 36, 745–758 (2018)

eller i form av ökad livskvalitet. Om en behandling ökar sannolikheten för att patienten lever i tidpunkt t med dp_t procentenheter och patientens nytta i perioden beräknas till $U(H_t, C_t)$ och om nyttan vid perfekt hälsa ges av $U(H^*, C_t)$ ges QALY-vinsten i perioden av $QALY_t = \frac{U(H_t, C_t)}{U(H^*, C_t)} \cdot dp_t$. Om vi använder antagandet om (multiplikativ) separabilitet mellan hälsa och materiell konsumtion blir uttrycket

$$QALY_t = \frac{u(H_t)}{u(H^*)} \cdot dp_t. \quad (1)$$

Notera att hälsovinsten i form av den ökade sannolikheten för överlevnad (dp_t) multipliceras med "vikten" $u(H_t)/u(H^*)$ för att omvandlas till QALY. Denna vikt baseras på patientens värderingar. Den svarar på frågan vad värdet (nyttan) av ett års liv med hälsan H_t är i förhållande till ett års liv med perfekt hälsa.

Om en behandling ökar livskvaliteten genom att föra patienten från ett hälsotillstånd H_t till ett bättre H'_t och om sannolikheten att patienten är vid liv i perioden ges av p_t ges QALY-vinsten i perioden av $QALY_t = p_t \cdot \frac{U(H'_t, C_t) - U(H_t, C_t)}{U(H^*, C_t)}$. Om vi använder antagandet om (multiplikativ) separabilitet mellan hälsa och materiell konsumtion blir uttrycket

$$QALY_t = p_t \cdot \frac{u(H'_t) - u(H_t)}{u(H^*)}. \quad (2)$$

Även i detta fall sätts nyttan av hälsovinsten i relation till nyttan av ett år med perfekt hälsa.

4.2.3 Betalningsvilja

Om det saknas alternativ till en viss medicinsk behandling är den samhällsekonomiskt motiverad om kostnaden per vunnet QALY är lägre än betalningsviljan per kvalitetsjusterat levnadsår (se Appendix A). Vi skriver detta villkor:

$$\frac{\text{Kostnad}_t}{QALY_t} \leq WTP_t, \quad (3)$$

där

$$WTP_t \equiv \frac{U(H^*, C_t)}{U_C(\tilde{H}_t, \tilde{C}_t)}$$

är samhällets betalningsvilja för ett levnadsår med perfekt hälsa. Betalningsviljan är det samhällsekonomiska värdet, i kronor, av ett ytterligare levnadsår med perfekt hälsa.²⁶

Betalningsviljan finns redan idag fastställd och vi behöver alltså inte utreda den närmare här. Det kan dock vara av intresse att notera att betalningsviljan för QALY måste fastställas med beaktande av att läkemedel i huvudsak finansieras via skatten. Det innebär för det första att betalningsviljan ska baseras på den genomsnittlige skattebetalarens vär-

²⁶Nämnamnaren är den representativa *skattebetalarens* marginalnytta av konsumtion.

dering av ett levnadsår. Den genomsnittlige skattebetalarens betalningsvilja, eller snarare *förmåga*, är sannolikt högre än betalningsförmågan för den genomsnittlige mottagaren av sjukvård. För det andra ska betalningsviljan för QALY beakta skatternas distortionära effekter på ekonomin. Eftersom läkemedel betalas via skatten och eftersom skatter orsakar kostnader för medborgarna som överstiger skatteintäkterna måste den fulla kostnaden för medborgarna beaktas.

Om ett nytt läkemedel ska användas istället för en etablerad behandling (som uppfyller kravet ovan) måste den nya behandlingen dessutom föredras framför den gamla. Den nya behandlingen är således endast samhällsekonomiskt motiverad om

$$\frac{\text{Extra Kostnad}}{\text{Extra QALY}} \leq 1. \quad (4)$$

Även merkostnad och extra hälsovinst måste diskonteras.

4.2.4 Diskontering

Vi ska nu analysera en medicinsk behandling som ger upphov till både hälsovinster och kostnader över i olika tidsperioder. Mer precist ger behandlingen hälsovinster $QALY_t$ och kostnaderna $Kostnad_t$ per patient vid tidpunkterna $t = 1$ och $t = 2$. De som lever får därmed betala $dC_t = -\frac{1}{p_t} \cdot Kostnad_t$. Behandlingen är samhällsekonomiskt motiverad om summan av de diskonterade kostnaderna delat med summan av de diskonterade QALY-vinsterna understiger betalningsviljan (se Appendix A). Vi skriver detta:

$$\frac{\sum_{t=0}^T \delta_K^t \cdot Kostnad_t}{\sum_{t=0}^T \delta_Q^t \cdot QALY_t} \leq WTP_0, \quad (5)$$

där δ_K är en diskonteringsfaktor för kostnader och δ_Q är en diskonteringsfaktor för QALY. Vi ska nu precisera hur dessa diskonteringsfaktorer kan bestämmas.

5 Diskonteringsfaktorerna

5.1 Diskontering av kostnader

Empiriska studier visar att människor värderar vinster och förluster olika beroende på när de uppstår. Att få en viss nytthet, t.ex. bättre hälsa eller extra materiell konsumtion, kan vara mer värt om nyttheten levereras redan idag än om den kommer först i morgon. Om vi får ett läkemedel idag måste vi betala för det genom minskad konsumtion av andra nyttigheter. En del av betalningen sker kanske redan idag, en annan del först i morgon. Frågan är hur vi ska värdera avbetalningsplanen i sin helhet. Svaret beror på hur vi ska värdera morgondagens betalningar i förhållande till dem vi gör redan idag.

Vanligen värderar vi en betalning i morgon lägre än samma summa idag. Även om dessa värderingar är intuitiva så är de bakomliggande orsakerna mer komplicerade.

Vi använder ett illustrativt exempel, potatis. Att få en potatis idag kan vara mer värt än att få en potatis i morgon. Ett första skäl är vår "otålighet". Allt annat lika, föredrar vi glädjen (nyttan) av en potatis idag framför glädjen (nyttan) av en potatis i morgon. Denna omständighet brukar kallas "ren tidspreferens" i den ekonomiska litteraturen. Ett andra skäl är det "utjämningsmotiv" som följer av att ökad konsumtion leder till en mättnadskänsla. Vi får det oftast materiellt sett bättre över tiden. Säg att vi bara har råd att äta en potatis idag men att vi kommer att ha råd att äta 10 potatisar i morgon. Om vi då skulle få en extra potatis i morgon så skulle den ge oss ganska lite extra glädje (nytta). Vi är redan ganska mätta på potatis. Men om vi skulle få en extra potatis idag så skulle den ge oss mycket extra glädje (nytta). Vi är fortfarande hungriga.

5.1.1 Offentliga sektorns (statens) låneränta

Vad skulle det då vara värt för oss att idag få ett löfte om säg fyra extra potatisar någon viss dag i framtiden? Värdet idag mäter vi med potatisar idag. Frågan är alltså hur många extra potatisar idag som är lika mycket värt som fyra extra potatisar i framtiden? Svaret kanske är att det räcker med en potatis idag. Varför? Mättnadskänslan i framtiden innebär kanske att fyra extra potatisar i framtiden ger oss lika mycket extra glädje då som den extra glädje vi skulle få av att äta två potatisar idag. Till det kommer den rena tidspreferensen (otåligheten) som kanske innebär att vi nöjer oss med bara en enda potatis idag. Alltså, fyra potatisar i framtiden är lika mycket värt som en potatis idag. Det betyder att om vi ska räkna nuvärdet av framtidens potatisar så måste vi räkna antalet potatisar och multiplicera det med en fjärdedel. Vi säger att diskonteringsfaktorn för potatis är en fjärdedel.

Nästa fråga är om det går att sätta ett värde på den personliga diskonteringsfaktorn? Vi kan kanske själva avgöra det genom introspektion. Men faktum är att även andra kan observera vad den är. För att beskriva detta fortsätter vi med potatis. Vi tänker oss att man kan låna (eller låna ut) potatis på en marknad. Långivarna är kanske villiga att ge oss en potatis idag om vi lovar att ge dem fyra potatisar i morgon. De extra potatisarna vi måste betala i morgon är låneräntan. Orsaken till att vi måste betala en låneränta är bl.a. att de som lånar ut potatis annars hade kunnat sätta den och fått tillbaka fyra potatisar i morgon. Det är denna förlust vi måste kompensera dem för. Annars är de inte villiga att låna oss potatis. Säg nu att vi själva har en diskonteringsfaktor på en femtedel. Det betyder att vi tycker konsumtion av fem extra potatisar i morgon är lika mycket värt som en extra potatis idag. Om vi då lånar en potatis idag och betalar fyra i morgon så har vi fått det bättre. Orsaken är att varje potatis i morgon bara är värd en femtedels potatis idag. Vi får alltså en potatis idag och betalar tillbaka fyra i morgon. Men för oss är dessa bara värda fyra femtedelars potatis. Alltså en vinst på en femtedels potatis. I frånvaro

av kreditrestriktioner fortsätter vi att låna allt fler potatisar. Men för varje ytterligare potatis vi äter idag ökar mättnadskänslan idag. Och för varje ytterligare potatis vi avstår i morgon ökar hungern i morgon. Vi är nöjda när vår personliga värdering av potatisar i morgon (alltså vår diskonteringsfaktor) är lika med låneräntan på marknaden. Slutsatsen är alltså att det är fullt möjligt att observera vilken personlig diskonteringsfaktor vi har för potatis. Den är lika med låneräntan på marknaden.

Vår slutsats är att framtida betalningar ska diskonteras med den marknadsmässiga låneräntan. Eftersom olika låntagare får låna till olika räntor (bl.a. beroende på risken att de inte kommer att kunna betala tillbaka) är det den specifika låntagarens ränta som ska användas. Om det är staten som är låntagaren är det statens låneränta som ska användas. Den är låg. Vi får som kollektiv, genom våra offentliga organisationer, bättre villkor på lånemarknaden, än om vi skulle låna privat.²⁷

5.1.2 Marknadsimperfectioner

Orsaken till att vi förordar att den offentliga sektorns (statens) låneränta bör vara utgångspunkt för TLV:s diskontering är en förmodan om att låneräntan uttrycker den samhällsekonomiska alternativkostnaden för den offentliga sektorns lån. Om det dock finns imperfectioner på lånemarknaden gäller inte denna förmodan. Då finns det anledning att sätta kalkylräntan högre än låneräntan. Vi kan se två möjliga imperfectioner.

1. I den mån staten har marknadsmakt på lånemarknaden (alltså att man har monopolmakt) är räntan på ytterligare lån högre än på tidigare tagna lån. Kalkylräntan bör i detta fall vara statens marginalränta.
2. I den mån statliga lån tränger undan privata investeringar som har ett högre värde för samhället än investerarnas egna privatekonomiska värden så överstiger den samhällsekonomiska kostnaden för ökad statlig upplåning räntan. I detta fall bör samhället räkna med en högre kalkylränta än sin låneränta.

Vi känner dock inte till någon utredning som substantierar att dylika imperfectioner skulle finnas eller vara kvantitativt betydelsefulla.

5.1.3 Korrelation mellan betalningar och ekonomisk aktivitet

Slutsatsen ovan är att den offentliga sektorns betalningar för läkemedel ska diskonteras med den riskfria låneräntan. Man kan dock fråga sig om man inte bör lägga till en riskpremie, för att ta hänsyn till att läkemedelskostnaderna kan vara korrelerade med den aggregerade risken i ekonomin.

²⁷Vi använder mycket riktigt staten för stora lån, både för att finansiera gemensam infrastruktur men även privat utbildning. Det innebär inte att vi skulle kunna använda staten som bank för t.ex. privata konsumtionslån eller för lån till privata entreprenörer. Då skulle staten stå inför likartade osäkerheter och problem som en vanlig bank. Det skulle i sin tur innebära att statens låneränta även skulle bli likartad med den vanliga banker betalar.

Motsvarande fråga har aktualiserats i diskussionerna kring höghastighetståg. Där ligger fokus på diskontering av de samhällsekonomiska intäkterna (som uppstår långt fram i tiden) snarare än kostnaderna (som tas i närmre tid). De samhällsekonomiska intäkterna av snabba tågförbindelser är bland annat de tillskott av produktion och konsumtion som kan komma till stånd. Dessa tillskott är högre under högkonjunktur och lägre under lågkonjunktur. Eftersom värdet av ett konsumtionstillskott är lägre i tider när konsumtionen redan är hög än i tider när den är låg måste man ta hänsyn till korrelationen mellan tillskotten och konjunkturcykeln.

Åter till frågan om läkemedelskostnader. Den avgörande frågan är om betalningar för läkemedel bidrar till skillnaden i materiellt välstånd mellan goda och dåliga tider. Är läkemedelsbetalningarna korrelerade med aggregerad konsumtion? Antag att den offentliga sektorn accepterar att ett visst läkemedel ska användas för vissa patienter till ett visst pris. Frågan är då om de framtida betalningarna för detta läkemedel kommer att vara korrelerade med den aggregerade konsumtionen i ekonomin. Det som skulle tala för ett risktillägg är t.ex. om användningen av läkemedlet ökar i lågkonjunkturer och om betalningarna sker vid användning. De samhällsekonomiska kostnaderna för ett dylikt läkemedel är högre än ett läkemedel med samma prislapp men som tenderar att konsumeras mer under högkonjunkturer.

En viktig invändning mot detta resonemang är dock att den offentliga sektorn (åtminstone staten) genom ökad eller minskad belåning kan fördela skatteuttaget i ekonomin mellan olika tidsperioder. Även om det skulle finnas en korrelation mellan läkemedelsanvändning och den ekonomiska konjunkturen följer det inte att det finns en korrelation mellan betalningar och konjunktur.

5.1.4 Överensstämmelse med andra sektorer och “second best”

Den diskontering vi förordar att TLV (och regionerna) ska använda vid värdering av läkemedel är sannolikt lägre än den som andra offentliga aktörer använder i andra sammanhang, t.ex. vid investeringar i vägar och tågförbindelser. Den kan även skilja sig från hur andra offentliga aktörer diskonterar kostnader i samband med andra investeringar i sjukvård. (Att den bara sannolikt är lägre beror på att vi inte vet vad en eventuell korrigering av statens låneränta med hänsyn till imperfektioner på kapitalmarknaderna skulle leda till.) Frågan är om dessa skillnader i sig leder till några problem, och om TLV av denna anledning bör använda samma diskontering som andra aktörer.

Den första frågan är då varför kalkylräntan är högre i andra sektorer. Vi ser åtminstone fyra möjliga förklaringar.

1. För det första kan det faktum att man använder en högre kalkylränta i andra sektorer än staten låneränta delvis bero på att man där lyckats korrigera för marknadsimperfektioner. Om så är fallet bör samma förfarande användas inom hälso-

och sjukvården.

2. För det andra handlar diskontering av infrastrukturinvesteringar i första hand om diskontering av samhällsekonomiska intäkter. Eftersom intäkterna är pro-cykliska måste de diskonteras med en risk-premie. En dylik risk-premie är inte påkallad för läkemedelskostnader. Om detta är orsaken till skillnaden så är det som det ska. Skillnaderna leder till korrekta investeringsincitament.
3. För det tredje är hälso- och sjukvården bunden av människovärdesprincipen som, i vår tolkning, innebär att framtida generationer ska värderas lika med nuvarande. Om man inte har en dylik princip inom transportsektorn kommer man att diskontera framtida intäkter (och kostnader) med en högre diskonteringsfaktor. Om detta är orsak till skillnaderna bör man överväga om sjuk- och hälsovårdens människovärdesprincip bör utsträckas även till andra sektorer. Om det är god etik i en sektor, torde det vara god etik i andra. Då är det alltså de andra sektorerna som bör anpassa sina kalkylräntor.
4. För det fjärde är det möjligt att den diskontering vi föreslår är den samhällsekonomisk korrekta och att andra offentliga aktörer gör fel.

Även om andra offentliga aktörer använder fel diskontering uppkommer frågan om inte TLV då ändå borde använda samma diskontering som de andra? Orsaken skulle vara att det annars kan finnas en risk för snedvridningar mellan olika offentliga verksamheter t.ex. inom sjukvården.

Om två verksamheter A och B har separata budgetar så spelar den ena verksamhetens val av diskonteringsfaktor inte någon roll för den andra verksamheten. Den diskonteringsfaktor som används inom verksamhet B påverkar inte B:s totala kostnad och därmed inte hur mycket som finns kvar till verksamhet A. Verksamhet A kan optimera inom sin budget genom val av rätt diskonteringsfaktor. Det är bara om verksamhet A och B delar en totalbudget som det kan bli problem.

För att svara på frågan måste vi först förtydliga att diskonteringsfaktorerna fyller två olika funktioner. Den primära funktionen är att välja rätt mellan olika läkemedelsprojekt med olika tidsprofiler för kostnader och intäkter. En för låg diskonteringsfaktor leder till att dåliga "dagsprojekt" (alltså projekt med stora nettovinster i närtid) tränger undan goda "framtidprojekt" (alltså projekt med stora nettovinster i framtiden). Problemet är att diskonteringsfaktorn även, om än på ett mer indirekt sätt, kommer att påverka den totala volymen projekt. Om projektkostnader i normalfallet uppträder tidigt och intäkterna uppkommer senare, kommer en beslutsfattare som använder en för låg diskonteringsfaktor (hög kalkylränta) att genomföra för få projekt inom dennes område.

Låt säga att så är fallet, att andra offentliga aktörer genomför för få projekt till följd av att de använder en för låg diskonteringsfaktor (hög ränta). Det leder i så fall till att

det finns för mycket budgetmedel kvar för läkemedel och att dessa därför överanvänds.²⁸ Frågan är då hur man ska minska denna överdrivna kostnad för läkemedel. Ett sätt är att diskontera läkemedel hårdare, mer precist, att diskontera läkemedel med samma låga diskonteringsfaktor som andra offentliga aktörer. Ett annat sätt är att sänka betalningsviljan för QALY för läkemedel. Det senare förfarandet är att föredra. Det leder till minskade kostnader för läkemedel, men utan att snedvrída valet av läkemedel med olika tidsprofiler.

Vi kan däremot inte avgöra hur stor sänkning av betalningsviljan för läkemedels-QALY skulle behöva vara.

5.2 Diskontering av QALY

Människors värderingar av nyttigheter varierar beroende på när vi får dem. Det gäller även nyttigheter som inte köps och säljs på en marknad, t.ex. hälsa och klimat. Framtida hälsovinster, mätta i QALY, måste således diskonteras för att vara jämförbara med hälsovinster idag. Offentliga beslutsfattare kan dock inte förutsätta att diskonteringsfaktorn för hälsa (och andra nyttigheter som medborgarna själva inte kan köpa och sälja på marknader) är lika hög som den som används för marknadsvaror.

Vi skiljer mellan två typer av hälsovinster, nämligen förbättrad livskvalitet och ökad sannolikhet för överlevnad. Dessa sammanfattas i hälsoekonomiska beräkningar i kvalitetsjusterade levnadsår, QALY. Hälsovinsterna räknas om till QALY år för år och diskonteras därefter. Vår analys här visar att diskonteringsfaktorn för QALY består av två komponenter. Mer precist ges diskonteringsfaktorn för QALY av

$$\delta_Q = \delta \cdot v.$$

Den första komponenten, δ , är den representativa medborgarens rena tidspreferens.

Den andra komponenten, v , beskriver hur värdet av ett års liv med perfekt hälsa förändras till följd av att den materiella konsumtionen förändras över tiden. En Taylorapproximation av uttrycket visar att

$$v = \frac{U(H^*, C_{t+1})}{U(H^*, C_t)} \approx \frac{U(H^*, C_t) + U_C(H^*, C_t) \cdot (C_{t+1} - C_t)}{U(H^*, C_t)} \approx 1 + \frac{C_{t+1} - C_t}{WTP_t},$$

där WTP_t är betalningsviljan för ett år med perfekt hälsa.²⁹ Om man mycket grovt räknar att BNP per capita är ca 500 tkr och att tillväxten är 2% per år så ökar den materiella konsumtionen med $C_{t+1} - C_t \approx 10$ tkr per år. Om betalningsviljan per QALY är ca 1 miljoner kronor, får vi att $v \approx 1,01$. Om vår gissning att den rena tidspreferensen

²⁸Om andra offentliga aktörer genomför för få projekt blir skuggpriset på skattekrönor för lågt. Ett för lågt skuggpris på skattekrönor innebär att betalningsviljan för QALY kommer att sättas för högt. Det innebär i sin tur att TLV's beslut kommer att leda till för höga offentliga utgifter för läkemedel.

²⁹Vi bortser här ifrån eventuella skillnader i marginalhyttan av konsumtion mellan patienter och skattebetalare.

ges av $\delta = 0,99$, blir konsekvensen att de två komponenterna tar ut varandra och att nettoeffekten således är att QALY inte ska diskonteras.

6 Diskontering med anledning av osäkerhet

Diskussionen ovan förutsatte att en behandling leder till säkra hälsoeffekter. Om en behandling endast leder till en viss hälsovinst med en viss sannolikhet måste kalkylen justeras. Vi kommer här att behandla det fall då man känner till sannolikheten för olika utfall. (Härledningarna återfinns i Appendix A.4.)

Hälsovinster Vad gäller hälsovinster behöver man beräkna förväntade QALY i varje tidsperiod och därefter diskontera dessa. Notera att QALY-måttet i sig tar hänsyn till att nyttan av ökad livskvalitet (t.ex. ökad kondition eller rörlighet) är avtagande. Därmed behöver man inte justera måttet för patienternas aversion mot hälsorisker.

Kostnader Att effekten är osäker reser frågan om huruvida samhällets betalning för läkemedlet bör betingas på effekten. När sannolikheterna är kända beror svaret på två omständigheter:

1. Avser risken läkemedlets effekt på enskilda patienter eller avser det effekt på populationsnivå?
2. Hur stor andel av befolkningen använder läkemedlet?

För beprövade läkemedel känner man till sannolikheten att det har effekt på patientnivå och det råder ingen aggregerad osäkerhet. Om läkemedlet används av en stor andel av befolkningen och om man vet att det fungerar i 50% av alla fall så spelar det inte någon roll om man betingar betalningarna på utfallet i det enskilda fallet. På grund av det stora antalet användare så kommer det knappt att bli någon variation i betalningarna.

Om läkemedlet endast används av en liten andel av befolkningen och om man vet att det fungerar i 50% av fallen så spelare det inte heller någon roll om man betingar betalningen på utfallet i de enskilda fallen. I detta fall kommer det uppstå stor variation i betalningarna, men eftersom betalningen delas upp på alla skattebetalare så har variationen endast en försumbar betydelse för var och en.

För nya läkemedel är osäkerheten framförallt aggregerad. Enkelt uttryckt så fungerar det (säg på alla) eller inte (säg på ingen). Om läkemedlet endast används av en liten andel av befolkningen spelar det återigen inte någon roll om betalningen betingas på utfallet. För det enskilde skattebetalaren är risken försumbar eftersom beloppen är så små.

I det fall osäkerheten är aggregerad och det används av en stor andel av befolkningen så måste betalningen risk-korrigeras på ett sätt som liknar risk-korrigeringen av hälsovinster. Skillnaden är att aversionen mot risker i materiell konsumtion torde vara en

annan än (mindre) än aversionen mot risker i hälsa och att risken måste justeras för hur stor andel av befolkningen som använder läkemedlet.

Risk-korrigeringen innebär att samhället föredrar att inte betinga betalningen på utfall. Dock kan det vara till fördel för läkemedelsföretag med nya läkemedel som är förknippade med aggregerad risk om staten tar på sig en del av risken. Detta ska dock betraktas som att företagen köper en tjänst av samhället för vilken man bör ersätta staten med minst den kostnad som risken innebär för skattebetalarna.

Notera att denna analys endast gäller i det fall sannolikheterna för olika utfall är kända av företagen men även av samhället.

Okända sannolikheter I de fall det är omöjligt att uppskatta sannolikheten för en viss hälsovinst bör värderingarna baseras på en försiktighetsprincip. Enkel uttryckt bör företagen inte ersättas för hälsovinster med okända sannolikheter, om man önskar betalning innan resultaten realiseras. Annorlunda uttryckt menar vi att man vid beräkning av förväntad hälsovinst ska använda de *lägsta* sannolikheter för goda utfall som är konsistenta med tillgänglig information.

Tyvärr hittar vi inte någon vägledning i den ekonomiska litteraturen för hur denna försiktighetsprincip skulle kunna operationaliseras. Ett motiv för försiktighetsprincipen är att människor hyser så-kallad osäkerhetsaversion. Detta finns belagt genom ekonomiska experiment (Ellsberg-paradoxen). Däremot finns det inte mycket vägledning i form av lämpliga beslutsmodeller.³⁰ Ett annat motiv för försiktighetsprincipen är att skapa incitament för företagen att ta fram och redovisa all information de har rörande läkemedlens hälsoeffekter.³¹

Om man kan finna former för en försiktighetsprincip vid värdering av läkemedel så skulle det i praktiken innebära att företagen får mycket starka incitament att konstruera betalningsmodeller där betalningarna skjuts på framtiden och betingas på resultaten. Detta förefaller även oss vara en lämplig lösning. Slutsatsen blir i så fall att det är viktigt att se till att det finns en kapacitet inom sjukvården att just arbeta med dylika betalningsmodeller.

³⁰Ett försök till formalisering är idén om så-kallad "max-min förväntad nytta" som innebär att man väljer det alternativ vars *värsta* möjliga utfall är så bra som möjligt. En mindre extrem formalisering bygger på så-kallad "Choquet förväntad nytta" och särskilt så-kallade "enkla kapaciteter". Det senare alternativet är dock långt ifrån operationaliserat i form av färdiga beslutsmodeller.

³¹En fullständig analys skulle baseras på så kallad "mechanism design". Vi känner tyvärr inte till några existerande arbeten som skulle ge ytterligare vägledning för hur dessa idéer skulle tillämpas i samband med prissättning av läkemedel.

A QALY-beräkningar och förväntad nyttoteori

A.1 Qaly och betalningsvilja i en period

Vi inleder framställningen med att härleda ett välkänt villkor för när en medicinsk behandling är samhällsekonomiskt motiverad. Villkoret är att kostnad per QALY understiger betalningsviljan för ett QALY. Vi härleder detta villkor i enklast möjliga sammanhang.

Låt säga att befolkningen består av N skattebetalare, varav N^S får behandlingen. Den representativa patientens förväntade nytta under ett visst år t ges av

$$EU_t = p_t \cdot U(H_t, C_t)$$

där p_t anger sannolikheten för att individen får leva under året, H_t är hälsotillståndet, C_t representerar konsumtion av varor och $U(H_t, C_t)$ beskriver nyttan under året. Enklast är att tänka att konsumtion av andra varor mäts i kronor. Vi förutsätter att nyttan ökar i båda nyttigheterna, om än i avtagande takt. Den representativa skattebetalaren, som inte får behandlingen, har nyttan

$$E\tilde{U}_t = U(\tilde{H}_t, \tilde{C}_t).$$

Samhällets sociala välfärdsfunktion är en vägd summa av de båda gruppernas förväntade nyttor.

Låt säga att en behandling ger upphov till kostnaden $Kostnad_t$ per behandlad person. Om hälsovinsten utgörs av en ökad chans för överlevnad, dp_t , så ges QALY av

$$QALY_t = \frac{U(H_t, C_t)}{U(H_t^*, C_t)} \cdot dp_t.$$

Om hälsovinsten istället utgörs av ökad livskvalitet dH_t så ges QALY av

$$QALY_t = \frac{U(H_t + dH_t, C_t) - U(H_t, C_t)}{U(H_t^*, C_t)} \cdot p_t.$$

Notera att samhällets betalningsvilja för ett år med perfekt hälsa ges av

$$WTP_t \equiv \frac{U(H_t^*, C_t)}{U_C(\tilde{H}_t, \tilde{C}_t)}.$$

Vi kan då visa att en behandling är samhällsekonomiskt motiverad om:

$$\frac{Kostnad_t}{QALY_t} \leq WTP_t,$$

förutsatt att $\frac{N^S}{N} \approx 0$. Vi härleder uttrycket separat för de två olika typerna av hälsovinster.

Ökad överlevnad Frågan är om en viss behandling som ökar sannolikheten för överlevnad med dp_t och som är behäftad med en viss kostnad $Kostnad_t$ ökar den aggregerade nyttan och därmed är samhällsekonomiskt motiverad.

Låt dp_t vara den ökade sannolikheten för överlevnad och $-dC_t$ vara den ändrade konsumtionen. Då får den representativa patienten nyttotillskottet

$$\Delta EU_t = (p_t + dp_t) \cdot U(H_t, C_t - dC_t) - p_t \cdot U(H_t, C_t)$$

d.v.s.

$$\Delta EU_t = U(H_t, C_t - dC_t) \cdot dp_t + p_t \cdot [U(H_t, C_t - dC_t) - U(H_t, C_t)]$$

där den första termen representerar hälsovinsten och den andra termen kostnaden. En Taylor-approximation av den andra termen ger

$$\Delta EU_t \approx U(H_t, C_t - dC_t) \cdot dp_t - p_t \cdot U_C(H_t, C_t) \cdot dC_t.$$

Den representativa skattebetalaren, som inte får behandlingen, får ett nyttotillskott om

$$\Delta E\tilde{U}_t \approx -U_C(\tilde{H}_t, \tilde{C}_t) \cdot dC_t.$$

Således är behandlingen samhällsekonomiskt motiverad om

$$dW = N^S \cdot [U(H_t, C_t - dC_t) \cdot dp_t - p_t \cdot U_C(H_t, C_t) \cdot dC_t] - (N - N^S) \cdot U_C(\tilde{H}_t, \tilde{C}_t) \cdot dC_t$$

eller

$$\frac{dW}{N^S} = U(H_t, C_t - dC_t) \cdot dp_t - \left[p_t \cdot U_C(H_t, C_t) + \frac{N}{N^S} \cdot \frac{N - N^S}{N} \cdot U_C(\tilde{H}_t, \tilde{C}_t) \right] \cdot dC_t$$

Notera att $dC_t = \frac{N^S \cdot Kostnad_t}{(N - N^S) + N^S \cdot (p_t + dp_t)} = \frac{\frac{N^S}{N}}{1 - \frac{N^S}{N} \cdot (1 - p_t - dp_t)} \cdot Kostnad_t$. Således

$$\frac{dW}{N^S} = U(H_t, C_t - dC_t) \cdot dp_t - \left[\frac{N^S}{N} \cdot p_t \cdot U_C(H_t, C_t) + \frac{N - N^S}{N} \cdot U_C(\tilde{H}_t, \tilde{C}_t) \right] \cdot \frac{Kostnad_t}{1 - \frac{N^S}{N} \cdot (1 - p_t - dp_t)}$$

Om $\frac{N^S}{N} \approx 0$ gäller att

$$\frac{dW/U_C(\tilde{H}_t, \tilde{C}_t)}{N^S} = \frac{U(H^*, C_t)}{U_C(\tilde{H}_t, \tilde{C}_t)} \cdot \frac{U(H_t, C_t)}{U(H^*, C_t)} \cdot dp_t - Kostnad_t.$$

Behandlingen är samhällsekonomiskt motiverad om

$$\frac{dW/U_C(\tilde{H}_t, \tilde{C}_t)}{N^S} = \frac{U(H^*, C_t)}{U_C(\tilde{H}_t, \tilde{C}_t)} \cdot \frac{U(H_t, C_t)}{U(H^*, C_t)} \cdot dp_t - Kostnad_t \geq 0.$$

Notera att samhällets betalningsvilja för ett år med perfekt hälsa ges av

$$WTP_t \equiv \frac{U(H^*, C_t)}{U_C(\tilde{H}_t, \tilde{C}_t)},$$

och att hälsovinsten av den ökade överlevnaden kan räknas om till QALY genom

$$QALY_t = \frac{U(H_t, C_t)}{U(H^*, C_t)} \cdot dp_t.$$

Således är behandlingen samhällsekonomiskt motiverad om:

$$\frac{Kostnad_t}{QALY_t} \leq WTP_t.$$

Ökad livskvalitet Frågan är om en viss behandling som ökar livskvaliteten med dH_t och som är behäftad med en viss kostnad $Kostnad_t$ ökar den aggregerade nyttan och därmed är samhällsekonomiskt motiverad. Den behandlade individen får nyttotillskottet

$$\Delta EU_t = p_t \cdot [U(H_t + dH_t, C_t - dC_t) - U(H_t, C_t)]$$

vilket kan dekomponeras i en hälsoeffekt och en kostnadseffekt

$$\Delta EU_t = p_t \cdot [[U(H_t + dH_t, C_t - dC_t) - U(H_t, C_t - dC_t)] + [U(H_t, C_t - dC_t) - U(H_t, C_t)]] .$$

En Taylor-approximation

$$\Delta EU_t \approx p_t \cdot [U(H_t + dH_t, C_t - dC_t) - U(H_t, C_t - dC_t)] - p_t \cdot U_C(H_t, C_t) \cdot dC_t.$$

$$\Delta EU_t \approx U(H_t, C_t - dC_t) \cdot dp_t - p_t \cdot U_C(H_t, C_t) \cdot dC_t.$$

Den representative skattebetalaren, som inte får behandlingen, får ett nyttotillskott om

$$\Delta E\tilde{U}_t \approx -U_C(\tilde{H}_t, \tilde{C}_t) \cdot dC_t.$$

Motsvarande beräkningar som ovan visar att behandlingen är samhällsekonomiskt om

$$\frac{dW/U_C(\tilde{H}_t, \tilde{C}_t)}{N^S} = \frac{U(H^*, C_t)}{U_C(\tilde{H}_t, \tilde{C}_t)} \cdot \frac{U(H_t + dH_t, C_t) - U(H_t, C_t)}{U(H^*, C_t)} \cdot p_t - \text{Kostnad}_t \geq 0.$$

Notera att samhällets betalningsvilja för ett år med perfekt hälsa ges av

$$WTP_t \equiv \frac{U(H^*, C_t)}{U_C(\tilde{H}_t, \tilde{C}_t)},$$

och att hälsovinsten av den ökade överlevnaden kan räknas om till QALY genom

$$QALY_t = \frac{U(H_t + dH_t, C_t) - U(H_t, C_t)}{U(H^*, C_t)} \cdot p_t.$$

Således är behandlingen samhällsekonomiskt motiverad om:

$$\frac{\text{Kostnad}_t}{QALY_t} \leq WTP_t.$$

A.2 Diskontering

För att beskriva diskontering väljer vi att endast härleda uttrycket för en behandling vars hälsovinst är ökad överlevnad. Fallet med ökad livskvalitet kan behandlas på samma sätt. Den representativa patientens förväntade nytta ges av

$$EU = \sum_{t=0}^T p_t \cdot \delta^t \cdot U(H_t, C_t).$$

En medicinsk behandling ger upphov till hälsovinster $dp_t > 0$ och kostnaderna Kostnad_t per patient. Förväntad nytta ges av

$$EU = \sum_{t=0}^T (p_t + dp_t) \cdot \delta^t \cdot U(H_t, C_t - dC_t).$$

Nyttotillskottet ges av

$$\Delta EU = \sum_{t=0}^T \delta^t \cdot [dp_t \cdot U(H_t, C_t - dC_t) + p_t \cdot (U(H_t, C_t - dC_t) - U(H_t, C_t))].$$

Där den första termen inom hakparantes anger hälsovinst och den andra termen anger kostnadseffekt. En Taylor approximation av kostnadseffekter ger

$$\Delta EU \approx \sum_{t=0}^T \delta^t \cdot [U(H_t, C_t - dC_t) \cdot dp_t - p_t \cdot U_C(H_t, C_t) \cdot dC_t].$$

På motsvarande sätt blir nyttotillskottet för en representativ skattebetalare, som inte behandlas

$$\Delta E\tilde{U} \approx - \sum_{t=0}^T \delta^t \cdot U_C(\tilde{H}_t, \tilde{C}_t) \cdot dC_t.$$

Notera att vi inte räknar med någon dödsrisk för dessa eftersom det tillkommer nya generationer. Behandlingen samhällsekonomiskt motiverad om

$$\frac{\Delta W}{N^S} = \Delta EU + \frac{N - N^S}{N^S} \cdot \Delta E\tilde{U} \geq 0.$$

Notera att $dC_t = \frac{N^S \cdot \text{Kostnad}_t}{(N - N^S) + N^S \cdot (p_t + dp_t)} = \frac{\frac{N^S}{N}}{1 - \frac{N^S}{N} \cdot (1 - p_t - dp_t)} \cdot \text{Kostnad}_t$ och om $\frac{N^S}{N} \approx 0$ gäller att

$$\Delta EU \approx \sum_{t=0}^T \delta^t \cdot U(H_t, C_t) \cdot dp_t.$$

$$\frac{N - N^S}{N^S} \cdot \Delta E\tilde{U} \approx - \sum_{t=0}^T \delta^t \cdot U_C(\tilde{H}_t, \tilde{C}_t) \cdot \text{Kostnad}_t$$

$$\frac{\Delta W}{N^S} = \sum_{t=0}^T \delta^t \cdot U(H_t, C_t) \cdot dp_t - \sum_{t=0}^T \delta^t \cdot U_C(\tilde{H}_t, \tilde{C}_t) \cdot \text{Kostnad}_t \geq 0.$$

QALY definieras av

$$QALY_t = \frac{U(H_t, C_t)}{U(H_t^*, C_t)} \cdot dp_t$$

och samhällets betalningsvilja för ett år med perfekt hälsa ges av

$$WTP_t \equiv \frac{U(H_t^*, C_t)}{U_C(\tilde{H}_t, \tilde{C}_t)}.$$

Således

$$\frac{\sum_{t=0}^T \delta^t \cdot \frac{U_C(\tilde{H}_t, \tilde{C}_t)}{U_C(\tilde{H}_0, \tilde{C}_0)} \cdot \text{Kostnad}_t}{\sum_{t=0}^T \delta^t \cdot \frac{U(H_t^*, C_t)}{U(H_0^*, C_0)} \cdot QALY_t} \leq WTP_0.$$

Notera att vi även kan skriva

$$\frac{U_C(\tilde{H}_t, \tilde{C}_t)}{U_C(\tilde{H}_0, \tilde{C}_0)} = \frac{U_C(\tilde{H}_1, \tilde{C}_1)}{U_C(\tilde{H}_0, \tilde{C}_0)} \cdot \frac{U_C(\tilde{H}_2, \tilde{C}_2)}{U_C(\tilde{H}_1, \tilde{C}_1)} \cdot \dots \cdot \frac{U_C(\tilde{H}_t, \tilde{C}_t)}{U_C(\tilde{H}_{t-1}, \tilde{C}_{t-1})} = \prod_{s=1}^t \mu_s,$$

där $\mu_s = \frac{U_C(\tilde{H}_s, \tilde{C}_s)}{U_C(\tilde{H}_{s-1}, \tilde{C}_{s-1})}$. På motsvarande sätt gäller att

$$\frac{U(H_t^*, C_t)}{U(H_0^*, C_0)} = \prod_{s=1}^t v_s$$

där $v_s = \frac{U(H_s^*, C_s)}{U(H_{s-1}^*, C_{s-1})}$.

Om vi tänker oss att $\mu_s = \mu$ och $v_s = v$ i varje tidpunkt så får vi

$$\frac{\sum_{t=0}^T \delta^t \cdot \mu^t \cdot \text{Kostnad}_t}{\sum_{t=0}^T \delta^t \cdot v^t \cdot \text{QALY}_t} \leq \text{WTP}_0.$$

Vi noterar vidare att om perfekt hälsa inte förändras över tiden $H_s^* = H^*$ och om vi approximerar uttrycket med en första ordningens Taylor approximation

$$v_s = \frac{U(H^*, C_s)}{U(H^*, C_{s-1})} \approx \frac{U(H^*, C_{s-1}) + U_C(H^*, C_{s-1}) \cdot (C_s - C_{s-1})}{U(H^*, C_{s-1})} = 1 + \frac{C_s - C_{s-1}}{\text{WTP}_{s-1}}.$$

Om man mycket grovt räknar att BNP per capita är ca 500 tkr och att tillväxten är 2% per år så ökar den materiella konsumtionen med $C_s - C_{s-1} \approx 10$ tkr per år. Om betalningsviljan per QALY är ca 1 miljoner kronor, får vi att $v_s \approx 1,01$.

Om vi istället använder en andra ordningens Taylor approximation

$$v = \frac{U(H^*, C_s)}{U(H^*, C_{s-1})} \approx 1 + \frac{C_s - C_{s-1}}{\text{WTP}_s} \cdot \left[1 + \frac{1}{2} \cdot \frac{U_{CC}(H^*, C_{s-1}) \cdot C_{s-1}}{U_C(H^*, C_{s-1})} \cdot \left(\frac{C_s - C_{s-1}}{C_{s-1}} \right) \right].$$

Vi noterar att $\frac{U_{CC}(H^*, C_{s-1}) \cdot C_{s-1}}{U_C(H^*, C_{s-1})}$ är koefficienten för relativ risk-aversion. Empiriska skattningar antyder att ett rimligt värde är ca 2. Detta innebär tillsammans med övriga antaganden ovan att

$$v = 1 + \frac{1}{100} \cdot \left[1 - \frac{1}{2} \cdot 2 \cdot \left(\frac{10}{500} \right) \right],$$

vilket knappt ändrar slutsatsen ovan.

A.3 Upplåning och skattedistortationer

I analysen ovan bortsåg vi ifrån att offentliga utgifter för läkemedel förutsätter distortionär beskattning och att staten kan skjuta betalningar mellan tidsperioder genom lån. Låt oss säga att en behandling ger effekter och ska betalas under två år. De betalningar som ska erläggas är Kostnad_0 respektive Kostnad_1 . Då innebär en optimal beskattning att samhället löser följande maximeringsproblem

$$\max_{dC_0, dC_1} U(H_0, C_0 - dC_0) + \delta \cdot U(H_1, C_1 - dC_1)$$

under bivillkoret av de två periodernas budgetrestriktioner. Skatteintäkten idag är $(1 - \lambda_0) \cdot dC_0$ vilket leder till en minskad konsumtion dC_0 för medborgaren. Skillnaden är den distortionära beskattningen. Differensen mellan intäkter och betalningar måste finansieras med en ökad belåning,

$$b = \text{Kostnad}_0 - (1 - \lambda_0) \cdot dC_0.$$

I framtiden betalar den offentliga sektorn Kostnad_1 och även lånet med ränta $(1 + r) \cdot b$ där r är upplåningsräntan. Detta leder till ett nödvändigt skatteuttag på

$$(1 - \lambda_1) \cdot dC_1 = \text{Kostnad}_1 + (1 + r) \cdot b$$

där λ_1 representerar den distortionära beskattningen i framtiden. Den konsoliderade budgetrestriktionen kan då skrivas

$$dC_0 + \frac{1}{1 + r} \cdot \frac{1 - \lambda_1}{1 - \lambda_0} \cdot dC_1 = \frac{1}{1 - \lambda_0} \cdot \text{Kostnad}_0 + \frac{1}{1 + r} \cdot \frac{1}{1 - \lambda_0} \cdot \text{Kostnad}_1.$$

Optimal beskattning definieras av

$$U_C(H_0, C_0) = \mu$$

och

$$\delta \cdot U_C(H_1, C_1) = \mu \cdot \frac{1}{1 + r} \cdot \frac{1 - \lambda_1}{1 - \lambda_0}$$

där μ är en Lagrange-multiplikator. Således krävs att

$$\delta \cdot \frac{U_C(H_1, C_1)}{U_C(H_0, C_0)} = \frac{1}{1 + r} \cdot \frac{1 - \lambda_1}{1 - \lambda_0}.$$

Av detta följer att de diskonterade kostnaderna för den representativa medborgaren ges av

$$dC_0 + \delta \cdot \frac{U_C(H_1, C_1)}{U_C(H_0, C_0)} \cdot dC_1 = dC_0 + \frac{1}{1 + r} \cdot \frac{1 - \lambda_1}{1 - \lambda_0} \cdot dC_1 = \frac{1}{1 - \lambda_0} \cdot \left[\text{Kostnad}_0 + \frac{1}{1 + r} \cdot \text{Kostnad}_1 \right].$$

Det tidigare villkoret gäller då fortfarande med ändringen att betalningsviljan för QALY ges av

$$WTP = \frac{U(H_0, C_0)}{(1 - \lambda_0)^{-1} \cdot U_C(H_0, C_0)}$$

och att

$$\delta_K = \delta \cdot \frac{U_C(H_1, C_1)}{U_C(H_0, C_0)} = \frac{1}{1 + r}.$$

Slutsatsen är att kostnaderna för läkemedlen utgörs av de diskonterade betalningarna till läkemedelsföretaget, där diskontering sker till den offentliga sektorns låneränta. I högerledet återges samhällets betalningsvilja för hälsovinster. Betalningsviljan beror endast

på dagens distortioner i beskattningen. Vi förutsätter att man redan tagit hänsyn till detta när man formulerat en betalningsvilja per QALY.

A.4 Osäkerhet (risk) rörande läkemedlens långsiktiga effekter

Det råder idag stora osäkerheter om ATMP-läkemedlens långsiktiga effekter. Vi kommer här att analysera hur en dylika osäkerheter påverkar bedömningen av deras värde. På grund av osäkerheterna om läkemedlens effekter finns det goda skäl att betinga betalningarna för ATMP på realiserade resultat, vilket innebär att betalningarna måste skjutas upp. Detta reser ytterligare frågor om hur samhället ska värdera dessa betalningar.

Här analyserar vi det fall då man har en god uppfattning om sannolikheterna för olika utfall. Detta brukar benämnas risk. För enkelhets skull bortser vi här ifrån tidsdiskontering. Låt säga att befolkningen består av N skattebetalare, varav N^S får behandlingen.

Under ett år t har en representativ patient förväntad nytta $EU_t = p_t \cdot U(H_t, C_t)$ i frånvaro av en behandling, där p_t anger sannolikheten för att individen får leva under året, $U(H_t, C_t)$ beskriver nyttan under året (odiskonterad), H_t är hälsotillståndet och C_t representerar konsumtion av varor. Enklast är att tänka att konsumtion av andra varor mäts i kronor. Vi förutsätter att nyttan ökar i båda nyttigheterna, om än i avtagande takt.

Behandlingen leder till hälsotillstånd H_t^i med sannolikhet q_t^i . Behandlingen betalas med dC_t^i . Förväntad nytta med en behandling blir därmed

$$p_t \cdot \sum_{i=1}^I q_t^i \cdot U(H_t^i, C_t - dC_t^i)$$

och förväntad nyttoökning ges av

$$dEU_t = p_t \cdot \sum_{i=1}^I q_t^i \cdot [U(H_t^i, C_t - dC_t^i) - U(H_t, C_t)].$$

Nyttoökningen kan dekomponeras i en hälsoeffekt och en betalningseffekt

$$dEU_t = p_t \cdot \sum_{i=1}^I q_t^i \cdot [U(H_t^i, C_t) - U(H_t, C_t)] + p_t \cdot \sum_{i=1}^I q_t^i \cdot [U(H_t^i, C_t - dC_t^i) - U(H_t^i, C_t)].$$

Betalningstermen approximeras med en andra gradens Taylor approximation

$$dEU_t = p_t \cdot \sum_{i=1}^I q_t^i \cdot [U(H_t^i, C_t) - U(H_t, C_t)] - p_t \cdot \sum_{i=1}^I q_t^i \cdot U_C(H_t^i, C_t) \cdot \left[dC_t^i - \frac{1}{2} \cdot \frac{U_{CC}(H_t^i, C_t) \cdot C_t}{U_C(H_t^i, C_t)} \cdot \frac{(dC_t^i)^2}{C_t} \right].$$

Använd $dC_t^i = \frac{N^S}{N} \cdot K_t^i$ och (nu explicit) uttryckt per behandlad person

$$\frac{dEU_t}{N^S} = p_t \cdot \sum_{i=1}^I q_t^i \cdot [U(H_t^i, C_t) - U(H_t, C_t)] - p_t \cdot \sum_{i=1}^I q_t^i \cdot U_C(H_t^i, C_t) \cdot \left[\frac{N^S}{N} \cdot K_t^i - \left(\frac{N^S}{N} \right)^2 \cdot \frac{1}{2} \cdot \frac{U_{CC}(H_t^i, C_t) \cdot C_t}{U_C(H_t^i, C_t)} \right]$$

Om endast en liten andel av befolkningen behandlas $\frac{N^S}{N} \approx 0$

$$\frac{dEU_t}{N^S} = p_t \cdot \sum_{i=1}^I q_t^i \cdot [U(H_t^i, C_t) - U(H_t, C_t)].$$

På motsvarande sätt har en skattebetalare, som inte behandlas, nyttoändringen

$$d\widetilde{EU}_t = - \sum_{i=1}^I q_t^i \cdot U_C(\tilde{H}_t, \tilde{C}_t) \cdot \left[\frac{N^S}{N} \cdot K_t^i - \left(\frac{N^S}{N} \right)^2 \cdot \frac{1}{2} \cdot \frac{U_{CC}(\tilde{H}_t, \tilde{C}_t) \cdot \tilde{C}_t}{U_C(\tilde{H}_t, \tilde{C}_t)} \cdot \frac{(K_t^i)^2}{\tilde{C}_t} \right].$$

Summera över N skattebetalare och beräkna per behandlad person

$$\frac{N - N^S}{N^S} \cdot d\widetilde{EU}_t = -U_C(\tilde{H}_t, \tilde{C}_t) \cdot \sum_{i=1}^I q_t^i \cdot \left[\frac{N - N^S}{N} \cdot K_t^i - \frac{1}{2} \cdot \frac{N - N^S}{N} \cdot \left(\frac{N^S}{N} \right) \cdot \frac{U_{CC}(\tilde{H}_t, \tilde{C}_t) \cdot \tilde{C}_t}{U_C(\tilde{H}_t, \tilde{C}_t)} \cdot \frac{(K_t^i)^2}{\tilde{C}_t} \right].$$

Om endast en liten andel av befolkningen behandlas $\frac{N^S}{N} \approx 0$

$$\frac{N - N^S}{N^S} \cdot d\widetilde{EU}_t = -U_C(\tilde{H}_t, \tilde{C}_t) \cdot \sum_{i=1}^I q_t^i \cdot K_t^i.$$

Summera $\frac{dW}{N^S} = \frac{dEU_t}{N^S} + \frac{N - N^S}{N^S} \cdot d\widetilde{EU}_t$

$$\frac{dW/U_C(\tilde{H}_t, \tilde{C}_t)}{N^S} = p_t \cdot \frac{U(H^*, \tilde{C}_t)}{U_C(\tilde{H}_t, \tilde{C}_t)} \cdot \sum_{i=1}^I q_t^i \cdot \frac{[U(H_t^i, C_t) - U(H_t, C_t)]}{U(H^*, \tilde{C}_t)} - \sum_{i=1}^I q_t^i \cdot K_t^i$$

$$WTP_t \equiv \frac{U(H^*, \tilde{C}_t)}{U_C(\tilde{H}_t, \tilde{C}_t)}.$$

$$E\{Kostnad_t\} = \sum_{i=1}^I q_t^i \cdot K_t^i$$

$$E\{QALY_t\} = \sum_{i=1}^I q_t^i \cdot \frac{[U(H_t^i, C_t) - U(H_t, C_t)]}{U(H^*, \tilde{C}_t)}$$

Behandlingen är samhällsekonomiskt motiverad om

$$\frac{E\{Kostnad_t\}}{E\{QALY_t\}} \leq WTP_t.$$

A.5 Osäkerhet om framtida materiell konsumtion och osäkerhet om hälsoeffekter

Om det även råder osäkerhet om framtida materiell konsumtion gäller att

$$\frac{Kostnad_0 + \delta \cdot E\{\mu\} \cdot E\{Kostnad_1\}}{QALY_0 + \delta \cdot E\{v\} \cdot E\{QALY_1\}} \leq WTP_0,$$

om hälsovinster och betalningar för behandlingen är oberoende av den framtida konsumtionen. Om det råder osäkerhet om framtida materiell konsumtion, t.ex. på grund av osäker tillväxt eller konjunktursvängningar, blir diskonteringsfaktorn en stokastisk variabel.

I vår diskussion om osäkerhet begränsar vi oss till en modell med enbart två perioder för att öka tydligheten. En behandling är samhällsekonomiskt motiverad om

$$WTP_0 \cdot [QALY_0 + \delta_Q \cdot QALY_1] - [Kostnad_0 + \delta_K \cdot Kostnad_1] \geq 0.$$

Vi kommer här att använda antagandet om att nyttofunktionen är multiplikativt separerbar i hälsa och materiell konsumtion. Vi illustrerar med det fall hälsovinsten är ökad sannolikhet för överlevnad. Då gäller att

$$QALY_t = \frac{u(H_t)}{u(H^*)} \cdot dp_t.$$

Diskonteringsfaktorn för QALY ges av

$$\delta_Q = \delta \cdot v = \delta \cdot \frac{U(H^*, C_1)}{U(H^*, C_0)} = \delta \cdot \frac{v(C_1)}{v(C_0)}.$$

Diskonteringsfaktorn för kostnader ges av

$$\delta_K = \delta \cdot \mu = \delta \cdot \frac{U_C(\tilde{H}, \tilde{C}_1)}{U_C(\tilde{H}, \tilde{C}_0)} = \delta \cdot \frac{v'(\tilde{C}_1)}{v'(\tilde{C}_0)},$$

om den representativa skattebetalarens hälsa är densamma över tiden, $\tilde{H}_1 = \tilde{H}_0$. Således

$$WTP_0 \cdot \left[\frac{u(H_0)}{u(H^*)} \cdot dp_0 + \delta \cdot \frac{v(C_1)}{v(C_0)} \cdot \frac{u(H_1)}{u(H^*)} \cdot dp_1 \right] - \left[Kostnad_0 + \delta \cdot \frac{v'(\tilde{C}_1)}{v'(\tilde{C}_0)} \cdot Kostnad_1 \right] \geq 0.$$

Om vi har osäkerhet om C_1 , \tilde{C}_1 , H_1 och $Kostnad_1$ blir den förväntade nyttovinsten

$$WTP_0 \cdot \left[\frac{u(H_0)}{u(H^*)} \cdot dp_0 + \delta \cdot \frac{E\{v(C_1) \cdot u(H_1)\}}{v(C_0) \cdot u(H^*)} \cdot dp_1 \right] - \left[Kostnad_0 + \delta \cdot \frac{E\{v'(\tilde{C}_1) \cdot Kostnad_1\}}{v'(\tilde{C}_0)} \right] \geq 0.$$

Vi föreställer oss att osäkerheten om framtida materiell konsumtion, C_1 , och framtida hälsotillstånd, H_1 , är oberoende av varandra. Då gäller att $E\{v(C_1) \cdot u(H_1)\} = E\{v(C_1)\} \cdot E\{u(H_1)\}$. Vi föreställer oss även att osäkerheten om framtida materiell konsumtion, \tilde{C}_1 , och framtida kostnader, $Kostnad_1$, är oberoende av varandra. Då gäller att $E\{v'(\tilde{C}_1) \cdot Kostnad_1\} = E\{v'(\tilde{C}_1)\} \cdot E\{Kostnad_1\}$. Således,

$$WTP_0 \cdot \left[\frac{u(H_0)}{u(H^*)} \cdot dp_0 + \delta \cdot \frac{E\{v(C_1)\}}{v(C_0)} \cdot \frac{E\{u(H_1)\}}{u(H^*)} \cdot dp_1 \right] - \left[Kostnad_0 + \delta \cdot \frac{E\{v'(\tilde{C}_1)\}}{v'(\tilde{C}_0)} \cdot E\{Kostnad_1\} \right]$$

Vi kan skriva detta som

$$WTP_0 \cdot [QALY_0 + \delta \cdot E\{v\} \cdot E\{QALY_1\}] - [Kostnad_0 + \delta \cdot E\{\mu\} \cdot E\{Kostnad_1\}] \geq 0.$$

Med andra ord

$$\frac{Kostnad_0 + \delta \cdot E\{\mu\} \cdot E\{Kostnad_1\}}{QALY_0 + \delta \cdot E\{v\} \cdot E\{QALY_1\}} \leq WTP_0.$$

B Den rena tidspreferensen som risk-aversion

För att tolka den rena tidspreferensen kommer vi i detta avsnitt diskutera hur den representativa patienten värderar olika läkemedel vars effekt är att öka patientens livslängd utan att påverka patientens livskvalitet (hälsa).

Denna övning visar att den så kallade rena tidspreferensen som ofta tolkas som ett uttryck för otålighet istället kan tolkas som en aversion mot risk i livslängd. Följaktligen kan den rena tidspreferensen också skattas genom experiment som testar just denna riskaversion. Tidigare undersökningar antyder en mycket kraftig diskontering.

För att kunna värdera små förändringar i livslängd definierar vi patientens nyttofunktion i kontinuerlig tid. Om patienten lever i T år är dennes totala nytta lika med

$$W(T) = \int_0^T e^{-\rho \cdot t} \cdot u(h(t), c(t)) \cdot dt$$

där ρ uttrycker patientens rena tidspreferens för kontinuerlig tid, $u(h, c)$ är patientens nyttonivå i varje enskild tidpunkt, $c(t)$ är konsumtionen i tidpunkt t och $h(t)$ är hälsan i tidpunkt t . Vi förutsätter här att nyttan av att vara död är lika med noll. Vi bortser ifrån utjämningsmotiv genom att anta att patientens materiella välbefinnande och hälsa är desamma

varje år som personen är vid liv, det vill säga att $c(t) = c$ och $h(t) = h$.

Notera att en liten förändring i patientens livslängd ger en ökad nytta som är lika med

$$W'(T) = e^{-\rho T} \cdot u(h, c)$$

och att förändringen av förändringstakten ges av

$$W''(T) = -\rho \cdot e^{-\rho T} \cdot u(h, c).$$

Således kan vi beräkna Arrow-Pratts koefficient för patientens absoluta aversion mot risker i livslängd genom

$$\frac{W''(T)}{W'(T)} = -\rho.$$

Den rena tidspreferensen är således identisk med patientens aversion mot risker i livslängd.³²

Detta förhållande visar på ett tankeexperiment som kan användas för att uppskatta den rena tidspreferensen. Låt säga att patienten ställs inför ett val mellan två olika mediciner. Den ena ger patienten $T \geq 1$ års extra livslängd med säkerhet, alltså nyttan $W(T)$ med säkerhet. Den andra medicinen är osäker och ger patienten $T + 1$ års extra livslängd med sannolikheten p eller $T - 1$ års extra livslängd med sannolikheten $1 - p$. Den osäkra medicinen ger alltså förväntad nytta $p \cdot W(T + 1) + (1 - p) \cdot W(T - 1)$. Frågan är vilken sannolikhet p som patienten kräver för att den osäkra medicinen ska vara lika attraktiv som den säkra. Om patienten svarar med en viss sannolikhet p så kan man dra slutsatsen att dennes rena tidspreferens ges av³³

$$\delta = \frac{1 - p}{p}.$$

Annorlunda uttryckt om patienten svarar att det räcker med lika sannolikheter $p = 1/2$ då kan vi dra slutsatsen att $\delta = 1$. (I detta fall kan vi även dra slutsatsen att $\rho = 0$ och således $W''(T) = 0$.) Om patienten är riskavers och svarar att det krävs $p = 0,6$ så kan vi dra slutsatsen att $\delta = 2/3$.

C Etiska utgångspunkter

I bästa fall kan ATMP bota olika sjukdomar. En engångsbehandling idag kan ge upphov till hälsovinster på lång sikt. Det värde vi åsätter en botande behandling idag beror

³²I diskret tid ges den rena tidspreferensen av $\delta = e^{-\rho}$.

³³Vi har $W(T) = pW(T + 1) + (1 - p)W(T - 1)$. Notera vidare att vi kan skriva om $W(T) = -e^{-\rho T}$ genom att utföra integralen och därefter genomföra en positiv affin transformation. Efter substitution $e^{-\rho T} = pe^{-\rho(T+1)} + (1 - p)e^{-\rho(T-1)}$ och division med $W(T - 1)$ erhålles $e^{-\rho} = pe^{-\rho^2} + (1 - p)$ det vill säga $\delta = p\delta^2 + (1 - p)$. Denna ekvation har lösningen $\delta = \frac{1-p}{p}$.

därmed på hur vi värderar framtida hälsovinster i relation till hälsovinster idag. Om samhällets betalning för en behandling idag täcks med ett lån eller om betalningen skjuts upp till en senare tidpunkt kommer även framtida generationer att vara med och finansiera dessa hälsovinster. De kostnader vi räknar med för en behandling idag blir därmed beroende av hur vi värderar framtida betalningar i relation till betalningar idag. Dyliga värderingar av framtida för- och nackdelar i relation till motsvarande för- och nackdelar idag benämns diskontering i den ekonomiska litteraturen. Det är underförstått att framtida för- och nackdelar värderas lägre, alltså att de diskonteras.

Diskontering reser etiska frågor. Jag kommer här att redovisa två olika förhållningssätt som leder till olika slutsatser om hur samhället bör diskontera framtida hälsovinster och betalningar.

C.1 Hälso- och sjukvårdslagen

Hälso- och sjukvårdslagen inleds med en etisk plattform som stadgar att vården ska ges med respekt för alla människors lika värde och för den enskilda människans värdighet. Detta brukar benämnas människovärdesprincipen. Vidare ska den som har det största behovet av hälso- och sjukvård ges företräde till vården (behovs- och solidaritetsprincipen). Verksamheten ska även vara organiserad så att den främjar kostnadseffektivitet (kostnadseffektivitetsprincipen).

Dessa etiska principer är enligt vår uppfattning förenliga med de välfärdsteoretiska grunderna för samhällsekonomiska kostnads-intäktsanalyser. Det finns dock ett par kontroversiella tolkningsfrågor som jag behöver beröra. Dessa rör prioritering av hälsoinsatser mellan olika skeden i livsrytmen, eller uttryckt annorlunda, prioritering av hälsoinsatser mellan olika generationer.

Människovärdesprincipen består av två delar. För det första har alla människor samma värde, t.ex. oberoende av inkomst och kön. För det andra ska den enskilda människans värdighet respekteras. Enligt statens medicinsk-etiska råd³⁴ ligger betydelsen av värdighet nära det som man benämner integritet. Integritet innebär att en person har rätt att få sina värderingar, önsknings- och åsikter respekterade. Man påminner om att integritet inte alltid kan betyda ett personligt självbestämmande. Ett uppenbart exempel är att små barn inte har denna förmåga. I sådana situationer kan representanter för sjukvården fatta beslut i patientens ställe med patientens bästa för ögonen. Om en läkare fattar beslut för en patient med patientens eget bästa för ögonen så talar man om medicinsk paternalism. Medicinsk paternalism kan dock bli problematisk om besluten grundas på felaktiga föreställningar om vad som är bäst för patienten.

³⁴Smer (2008). Etik - En introduktion. Etiska vägmarken 1. Frizes, Stockholm.

Problemet Det problem som vi behöver konfrontera är att människovärdesprincipens två delar “människors samma värde” och “människans värdighet” kan komma att stå i motsatsställning till varandra, beroende på hur de tolkas. Mer precist gäller frågan om samhällets värdering av en hälsovinst ska påverkas av personens “kronologiska ålder”, alltså hur gammal personen är. För att ställa frågan på sin spets kan man fråga sig om det är lika mycket värt att rädda livet på en tjuugoåring som på en sextioåring, om det i båda fallen endast handlar om att förlänga livet med ett år?

I förarbetena till hälso- och sjukvårdslagen (proposition 1996/97:60) slår man fast att begåvning, social ställning, inkomst, ålder etc. inte får avgöra vem som skall få vård eller kvaliteten på vården. Man menar att det är en form av diskriminering och oförenligt med de etiska principerna att generellt låta behoven stå tillbaka på grund av ålder, födelsevikt, livsstil eller ekonomiska och sociala förhållanden. Däremot är det förenligt med de etiska principerna att i det enskilda fallet ta hänsyn till omständigheter som begränsar nyttan av medicinska åtgärder.

Enligt Lars Sandman och Emelie Heintz (2013) måste man göra en distinktion mellan kronologisk och biologisk ålder där biologisk ålder avgörs av organens och kroppens funktion. Sandman och Heintz menar att människovärdesprincipen inte medger att man tar hänsyn till kronologisk ålder hos patienten eller att man tillämpar generella kronologiska åldersgränser vid prioriteringar, om inte den kronologiska åldern generellt är förknippad med omständigheter som begränsar eller påverkar effekten av en behandling. En patients biologiska ålder i form av nedsatta fysiologiska resurser som bl.a. påverkar möjligheten att tillgodogöra sig åtgärder kan dock vägas in.

Såvitt jag förstår innebär detta att förarbetena bör tolkas på följande sätt:

1. Det är mer värt att rädda livet på en tjuugoåring än på en sextioåring, om den förre kan förväntas leva fler år än den senare. (Denna prioritering är baserad på biologisk ålder, och därför förenlig med människovärdesprincipen. Mer precist kan den härledas ur behovsprincipen och/eller kostnadseffektivitetsprincipen.)³⁵
2. Det är lika mycket värt att rädda livet på en tjuugoåring som på en sextioåring, om det i båda fallen handlar om att förlänga livet med ett år. Motsatsen hade varit en prioritering baserat på kronologisk ålder.

Ekonomisk forskning visar dock att människor själva lägger vikt vid kronologisk ålder. Man värderar bl.a. hälsoinsatser vid olika åldrar olika. Diskontering (i bemärkelsen ren tidspreferens) av framtida för- och nackdelar förefaller vara en närmast universell preferens. Enligt vår uppfattning synes därmed integritetsprincipen innebära att värdering av hälsovinster måste baseras på kronologisk ålder, inte endast biologisk ålder.

³⁵Denna fråga debatterades under våren 2020 på DN debatt av bl.a. Torbjörn Tännsjö (<https://www.dn.se/debatt/vi-bor-radda-de-unga-om-varden-inte-kan-klara-alla/>). Tännsjö förordar en dylik prioritering men menar att detta strider mot förarbetena. Det senare påståendet motsägs dock av Lars Sandman, Niklas Juth och Erik Gustavsson respektive Johnny Ludvigsson på delvis olika grunder.

C.2 Intertemporala val ur ett livcykelperspektiv

För att analysera samhällets problem med att fördela t.ex. sjukvårdsresurser mellan olika ändamål brukar man i den ekonomiska litteraturen tala om en så kallad social planerare. Denne representerar alla olika offentliga beslutsfattare som påverkar prioriteringarna inom sjukvården, från behandlande läkare via klinikchefer till riksdag och regering. Den sociale planeraren fördelar tillgängliga resurserna mellan olika personer i varje tidpunkt. Men eftersom samhället kan spara och låna resurser, och eftersom vissa medicinska ingrepp påverkar hälsan både omedelbart och i framtiden, måste man även tänka sig att den sociale planeraren gör en plan som fördelar resurser mellan de som lever i olika tidpunkter.

Vi börjar med att tänka oss att den sociale planeraren respekterar integritetsprincipen. Det brukar benämnas icke-paternalistiska preferenser i den ekonomiska litteraturen. Det betyder att den sociale planeraren endast bryr sig om medborgarnas välfärd. I valet mellan två medicinska behandlingar till samma individ väljer den sociale planeraren den som leder till högst nytta för just denne individ. I valet mellan två medicinska behandlingar som ges till två olika personer spelar det både roll hur dessa själva värderar sina respektive behandlingar och hur den sociale planeraren väger nyttovinster för en person mot nyttovinster för en annan.

För att beskriva detta tänker man sig att den sociale planeraren gör dessa val för att maximera en så kallad social välfärdsfunktion. En social välfärdsfunktion summerar och väger samman den nytta som medborgarna får ut av olika hälsoinsatser. Således tänker man på det samhällsliga värdet av hälsoinsatser i två steg. För det första leder en hälsoinsats till ökad nytta för den person som får insatsen. För det andra leder ökad nytta för en medborgare till ett högre värde på den sociala välfärdsfunktionen. Värdet av en hälsoinsats beror således på två olika värderingar, nämligen hur patienten värderar insatsen och på hur samhället värderar denna person.

Den sociala välfärdsfunktionen uttrycker med andra ord endast samhällets fördelningspolitiska preferenser, hur en nyttovinster för en person ska vägas mot en nyttovinster för en annan. Avvägningar mellan olika hälsovinster och andra nyttigheter för en och samma individ görs utifrån dennes egna värderingar.

Ett exempel Vi åskådliggör resonemanget med ett enkelt exempel. Låt oss säga att samhället består av två individer. Låt oss även säga att dessa värderar hälsa och konsumtion av andra varor på precis samma sätt. Vi beskriver detta med en nyttofunktion $\tilde{U}_i = U(H_i) + C_i$ där H_i är personen i 's hälsa och C_i är konsumtion av andra varor. Nyttan ökar med ökad hälsa men varje ytterligare hälsotillskott ger upphov till allt lägre nyttoavkastning (konkavitet). För enkelhets skull antar vi att nyttan av konsumtion av andra varor är linjär.

Vidare illustrerar vi resonemanget med en så kallad utilitaristisk social välfärdsfunk-

tion³⁶ som innebär att man summerar medborgarnas nyttor, med lika vikter, d.v.s.

$$W = \tilde{U}_1 + \tilde{U}_2 = [U(H_1) + C_1] + [U(H_2) + C_2].$$

Låt oss även säga att de två personerna kan ha olika hälsa, \tilde{H}_1 respektive \tilde{H}_2 , utan sjukvård. Samhället erbjuder sjukvårdsinsatserna h_1 och h_2 vilket innebär att hälsan blir $H_i = \tilde{H}_i + h_i$ för respektive person. Kostnaden för sjukvård är $c \cdot h_i$. Denna kostnad finansieras genom en skatt som innebär att medborgarnas konsumtionen av andra varor måste sänkas med samma belopp. Den sociala välfärdsfunktionen kan då skrivas

$$W = U(\tilde{H}_1 + h_1) + U(\tilde{H}_2 + h_2) + C_1 + C_2 - c \cdot h_1 - c \cdot h_2.$$

Den sociale planeraren väljer då sjukvårdsinsatserna

$$U_H(\tilde{H}_i + h_i) = c,$$

där subindex indikerar derivata. Således får båda de sjukvårdsinsatser som innebär att deras marginella nytta av sjukvård blir den samma. Detta innebär att medborgarna i detta fall i slutändan får samma hälsa

$$H_i = \tilde{H}_i + h_i = U_H^{-1}(c).$$

Annorlunda uttryckt kommer en person att få insatsen

$$h_i = U_H^{-1}(c) - \tilde{H}_i,$$

vilket innebär att den som har sämre hälsa i utgångsläget erbjuds mer sjukvård. Detta följer av att människor förutsätts ha avtagande nytta av hälsotillskott (konkavit) och att den sociala välfärdsfunktionen sätter lika stort värde för nyttotillskott oberoende av individ.

Avvägningen mellan olika sorters vård och andra nyttigheter En viktig poäng är att dessa sjukvårdsinsatser är precis de som var och en skulle välja själv. Det vill säga att

³⁶Mer allmänt kan man visa att den sociala välfärdsfunktionen under vissa antaganden måste ges av $W = \sum_i w(\tilde{U}_i)$ där \tilde{U}_i är individ i 's nytta och w är en växande funktion. De substantiella antaganden som ligger till grund för detta resultat är att den sociala välfärdsfunktionen ökar när någon medborgares nytta ökar, att alla medborgares nytta behandlas på samma sätt, samt att en persons nytta inte påverkar avvägningen mellan två andra personers nytta. Om man dessutom kräver att alla avvägningar ska vara oberoende av nyttans måttenhet (jämför att man kan mäta kostnader i kronor eller ören) så gäller att $w(U) = \frac{1}{\alpha} \cdot U^\alpha$. Parametern α beskriver samhällets preferenser för jämlikhet. Den sociala välfärdsfunktionen kallas utilitaristisk om $\alpha = 1$. Detta innebär att samhället inte bryr sig om skillnader i nytta mellan olika medborgare. Om man ogillar skillnader i nytta så gäller att $\alpha < 1$. I extremfallet $\alpha \rightarrow -\infty$ bryr man sig endast om den sämst ställda medborgaren. Se vidare Herve Moulin (2004). Fair Division and Collective Welfare. Cambridge, Massachusetts: MIT Press.

den sociale planeraren respekterar den avvägning mellan hälsa och och andra nyttigheter som medborgarna själva önskar göra. Och om vi skulle tänka oss att hälsan består av flera olika komponenter, t.ex. kondition och rörlighet, så kommer den sociale planeraren även att göra samma avvägning mellan insatser för förbättrad kondition och insatser för förbättrad rörlighet som medborgarna själva skulle vilja göra.

Avvägningen mellan olika personer Notera dock att om de båda individerna betalar lika mycket skatt så innebär prioriteringarna ovan att samhället omfördelar välstånd från den person som har bättre hälsa i utgångsläget till den som har sämre. Det är här den sociala välfärdsfunktionen kommer in. Den utilitaristiska välfärdsfunktionen innebär att samhället bejakar omfördelningar mellan olika personer om nyttoförlusten för den som vinner är större än nyttoförlusten för den som förlorar.

Även detta förfarande kan vi dock härleda tillbaka till medborgarnas egna preferenser, och därmed till integritetsprincipen. Innan en person vet om den kommer att ha tur och bli den med bättre hälsa i utgångsläget eller om den kommer att ha otur och bli den med sämre hälsa i utgångsläget så bör båda möjligheterna te sig lika sannolika. Mer allmänt bör sannolikheten för olika hälsotillstånd överensstämja med motsvarande befolkningsandelar. Om man således ber medborgarna att utforma ett hälso- och sjukvårdssystem “bakom okunskapens slöja” och om dessa då ser till sin *förväntade* nytta så kommer man precis som den utilitaristiska välfärdsfunktionen att lägga lika vikt vid båda möjligheterna och därmed föreskriva vårdinsatser baserat på behov.

Den utilitaristiska sociala välfärdsfunktionen intar en särställning bland sociala välfärdsfunktioner i det att den respekterar både människors lika värde och integritetsprincipen även i situationer när människor är osäkra om t.ex. sina framtida inkomster och hälsotillstånd.³⁷

Intertemporala val Vi använder motsvarande förfarande för att analysera intertemporala val. Låt säga att medborgarna lever i två perioder, som ung och gammal, och att alla har samma nyttofunktion som är additivt separerbar

$$\tilde{U} = U(H_1) + \delta \cdot U(H_2)$$

där H_1 är hälsan som ung och H_2 är hälsan som gammal. Vi tänker oss att hälsa kan mätas på samma sätt i båda perioderna. Ett exempel är kondition (maximal syreupptagningsförmåga). Vi tänker oss för tillfället även att konditionen påverkar personens välbefinnande på samma sätt som gammal och som ung. Detta antagande uttrycks genom att vi har samma nyttofunktion U i båda perioderna. (Vi kommer senare att släppa detta antagande.) Parametern δ är medborgarnas “rena” tidspreferens. Den rena tidspreferensen beskriver

³⁷Harsanyi

en persons relativa värdering av hälsovinster och andra nyttigheter som åtnjuts vid olika (kronologiska) åldrar. Vi kommer senare återkomma till hur stor denna är, men empiriska skattningar visar att människor värderar senare nyttigheter lägre än tidigare, det vill säga att $\delta < 1$.

Vi tänker oss för enkelhets skull att man genom olika sjukvårdsinsatser kan fördela hälsan fritt mellan olika tidpunkter, så länge $H_1 + H_2 = H$. Denna restriktion härrör sig bland annat från det faktum att vi har begränsade resurser. Om medborgaren själv får välja kommer denne att välja en bättre hälsa i den första perioden än i den andra, det vill säga $H_1 < H_2$ när $\delta < 1$.

Om hälsan i frånvaro av sjukvårdsinsatser är lägre vid hög ålder än vid unga år, kan personen mycket väl föredra mer sjukvårdsinsatser vid hög ålder än vid unga år. Detta är en fråga vi fördjupar oss i senare, men som vi bortser ifrån här. Här vill vi fokusera på frågan om medborgarnas rena tidspreferens bör få påverka samhällets värdering av hälsovårdsinsatser.³⁸

För att representera att samhället består av olika generationer bruka man använda en modell med överlappande generationer. Det vill säga i varje tidpunkt finns det några som är unga och några som är gamla. Vi låter då

$$\tilde{U}^t = U(H_1^t) + \delta \cdot U(H_2^t)$$

representera nyttan för en individ som är ung i tidpunkt t och följaktligen gammal i tidpunkt $t + 1$. Den utilitaristiska sociala välfärdsfunktionen summerar nyttan för alla individer, med lika vikter, från tidernas begynnelse till dess slut

$$W = \sum_t \tilde{U}^t = \sum_t [U(H_1^t) + \delta \cdot U(H_2^t)].$$

Den sociale planeraren maximerar denna funktion under bivillkoret av en resursrestriktion, t.ex. att resurser kan fördelas fritt mellan generationerna $\sum_t [H_1^t + H_2^t] = H$.³⁹ Det är lätt att se att de som föds i en viss tidpunkt t kommer att få en högre hälsa som unga än som gamla, det vill säga att $H_1^t > H_2^t$ om $\delta < 1$. Detta gäller naturligtvis för samtliga generationer t . Detta är ett uttryck för medborgarnas egna tidspreferenser.

Vår slutsats är att en social välfärdsfunktion som ger alla personer samma värde och som respekterar människors integritet, även avseende intertemporala val, innebär att

³⁸Att lägre hälsa i utgångsläget i äldre år innebär att människor önskar lägga en större andel av sjukvårdsresurserna till äldre år (allt annat lika) kan beskrivas som ett uttryck för diskontering. Vi uppfattar en dylik diskontering som såväl ett uttryck för människovärdesprincipen som behovsprincipen. Vi uppfattar tror dessutom inte heller att det råder delade meningar om huruvida detta överensstämmer med den etiska plattformen.

³⁹Denna formulering sopar ett matematiskt problem under mattan. Vi tänker oss att tiden är oändlig. Det betyder att summorna inte självklart är väldefinierade. Det finns två sätt att hantera problemet. Det ena är att vi tillåter en diskontering av framtida generationer. Det andra är att vi ersätter summan med medelvärdet, eller mer precist $W = \liminf_{T \rightarrow \infty} \frac{1}{2T} \sum_{t=-T}^T \tilde{U}^t$. Motsvarande gäller resursrestriktionen.

samhällets värdering av hälsovinster delvis baseras på människors tidspreferenser, det vill säga kronologisk ålder

C.3 Intertemporala val ur ett beslutsfattarperspektiv

Vi föreställer oss dock att många beslutsfattare i verkligheten, t.ex. läkare, klinikchefer och sjukvårdspolitiker, inte ser beslutsproblemet framför sig på det sätt som vi beskrivit ovan. Vi skissar här hur vi förmodar att problemet ser ut ifrån deras horisont. I brist på bättre ord kommer vi att benämna detta "beslutsfattarperspektivet". Den huvudsakliga skillnaden är att det vi ovan beskrev som prioriteringar mellan olika skeden under livscykeln för beslutsfattaren framstår som prioriteringar mellan olika generationer.

Beslutsfattarens mest påtagliga problem är att fördela hälsoinsatser mellan de generationer som lever idag (som vi kallar tidpunkt t). Beslutsfattaren har egna preferenser för detta problem som kan skrivas

$$V^t = U(H_1^t) + w \cdot U(H_2^{t-1}),$$

där w är beslutsfattarens egen vikt på den äldre generationen (i relation till vikten på den yngre som vi satt till 1). Men beslutsfattaren (möjligen på en högre administrativ nivå) ser naturligtvis även till avvägningar mellan hälsoinsatser idag och hälsoinsatser i morgon. Beslutsfattarens preferenser för detta problem kan skrivas

$$V = \sum_t \beta^t \cdot V^t$$

där $\beta \leq 1$ är beslutsfattarens egen diskonteringsfaktor för hälsovinster idag och i framtiden. Sammantaget innebär detta att beslutsfattarens preferenser kan skrivas

$$V = \sum_t \beta^t \cdot [U(H_1^t) + w \cdot U(H_2^{t-1})].$$

Vi tänker oss vidare att beslutsfattaren har samma resursrestriktion som tidigare, det vill säga $\sum_t [H_1^t + H_2^t] = H$.

Relation till människovärdesprincipen För att relatera hur intertemporala val görs i praktiken (så som vi beskrivit dem ovan) till människovärdesprincipen skriver vi om beslutsfattarens preferenser på följande sätt

$$V = \sum_t \beta^t \cdot [U(H_1^t) + w \cdot \beta \cdot U(H_2^t)].$$

Detta är exakt samma uttryck som det föregående, med skillnaden att vi här summerar över generationer och inte över tidpunkter. Kom ihåg att w representerar den vikt som

beslutsfattaren själv lägger vid den äldre generationen inom en viss tidpunkt och att β representerar beslutfattarens egen värdering av framtida (nytt-) vinster i förhållande till motsvarande vinster idag, oavsett vilken generation de tillfaller. Den utilitaristiska sociala välfärdsfunktion som uttrycker integritetsprincipen och likavärdesprincipen ges av

$$W = \sum_t [U(H_1^t) + \delta \cdot U(H_2^t)].$$

En jämförelse leder till två slutsater.

1. När vi fokuserar på avvägningen mellan hälsoinsatser vid unga år och äldre år, för en och samma person, så kommer beslutfattaren att göra samma avvägning som personen själv om och endast om $w \cdot \beta = \delta$. Detta är alltså ett villkor för att integritetsprincipen ska respekteras.
2. Alla generationer att behandlas lika om och endast om $\beta = 1$. Detta är alltså ett villkor för likavärdesprincipen.

Sammantaget innebär det att beslutfattarperspektivet är fullt förenligt med människovärdesprincipen, så som vi tolkar den. Dock ställer det krav på beslutfattarens diskonteringsfaktor och vikt vid olika generationer som lever i samma tidpunkt. Beslutsfattaren respekterar både integritetsprincipen och att olika generationer värderas lika om och endast om $\beta = 1$ och att beslutfattaren lägger vikten $w = \delta < 1$ vid den äldre generationen i varje tidpunkt.

Annorlunda uttryckt bör man betrakta den politiska diskonteringsfaktorn β i huvudsak som en värdering av olika generationer. I den ekonomiska litteraturen synes det finnas en konsensus om att denna diskonteringsfaktor bör sättas till ett (se Gollier ...), det vill säga att alla generationer ska värderas lika. Vi gör inte någon annan bedömning. Detta synes även vara en rimlig tolkning utifrån människovärdesprincipen som styr den svenska sjukvården.

Den mer kontroversiella slutsatsen här är att hälsovinster som tillfaller en äldre generation bör värderas lägre än motsvarande hälsovinster som tillfaller en yngre generation i samma tidpunkt. Annorlunda uttryckt: om man endast betraktar valen i varje enskild tidpunkt förefaller de vara diskriminerande.⁴⁰ Enligt vår uppfattning framstår dessa val som diskriminerande endast på grund av beslutfattarperspektivet.

C.4 Människovärdesprincipen enligt förarbetena

Enligt Proposition 1996/97:60 innebär människovärdesprincipen att en viss hälsovinster i en viss tidpunkt ska värderas på samma sätt oberoende av om den tillfaller en äldre eller

⁴⁰Dessutom, även om dessa val överensstämmer med de val varje person själv skulle vilja göra, finns det ett slags tidsinkonsistensproblem. I varje tidpunkt kan den äldre generationen vilja förespråka att $H_1^t = H_2^{t-1}$. Man har ju då redan fått fördelarna och kan vilja avstå från nackdelarna.

en yngre person. Det innebär att beslutsfattaren måste sätta $w = 1$.

Om vi även tolkar människovärdesprincipen så att hälsovinster för personer av en viss ålder som lever i olika tidpunkter ska värderas lika, alltså att $\beta = 1$, blir konsekvensen att $H_1^t = H_2^t$. Det innebär att de beslutsfattaren åsidosätter integritetsprincipen, när det gäller intertemporala val.⁴¹

Om man däremot tillåter beslutsfattarens diskonteringsfaktor att överensstämma med medborgarnas tidspreferenser, det vill säga att $\beta = \delta$, så kommer avvägningen av hälsovinster inom en generation vara de samma som medborgarna själv önskar. Med andra ord överensstämmer valen med integritetsprincipen. Konsekvensen är dock att framtida generationer ges en lägre vikt än dagens generationer. Detta strider dock mot människors lika värde, när det gäller personer som tillhör olika generationer.

Även olika mellanpositioner kan identifieras om man utgår ifrån propositionens tolkning av människovärdesprincipen, men accepterar att hälsovinster som uppträder i morgon får värderas lägre än hälsovinster som uppträder idag. Om beslutsfattarens diskonteringsfaktor sätts på en nivå mellan medborgarnas egen diskonteringsfaktor och ett (alltså ingen diskontering) uppstår en avvägning mellan likavärdesprincipen (mellan generationer) och integritetsprincipen (avseende intertemporala val).

Som vi ser det implicerar således den tolkning av människovärdesprincipen som ges i förarbetena ett etiskt dilemma. Det går inte att förena likavärdesprincipen och integritetsprincipen. Detta dilemma försvinner om människovärdesprincipen istället tolkas ur ett livcykelperspektiv, så som vi visade ovan.

⁴¹Att integritetsprincipen åsidosätta innebär även att utfallet inte är Pareto-optimalt, på följande sätt. I unga år, skulle varje person tacka ja till ett avtal som ger honom eller henne en gåva (en hälsovinst) från den då levande äldre generationen. I utbyte förbinder sig personen att i äldre år ge exakt samma gåva till en person som då är ung. Sådana avtal kan naturligtvis inte skrivas i praktiken, men samhället kan genom sin sjukvårdspolitik implementera samma utfall. För detta krävs dock att människovärdesprincipen ges en korrekt tolkning.