



“Det handlar om miljarder”

En metodanalys av
hur assistansfusket bedöms
av svenska myndigheter

Niklas Altermark & Hampus Nilsson

”Det handlar om miljarder”

En metodanalys av hur assistansfusket bedöms av svenska myndigheter
Niklas Altermark & Hampus Nilsson

Citat från engelskspråkiga publikationer är genomgående översatta av författarna till denna rapport.

Utgivare: STIL, Stiftarna av Independent Living i Sverige

ISBN 978-91-977929-3-6

Tryck: Exellent grafisk form, 2017

Form: Freke Råihä

”Det handlar om miljarder”

En metodanalys av
hur assistansfusket bedöms av
svenska myndigheter

Fil.dr. Niklas Altermark & Hampus Nilsson
Statsvetenskapliga institutionen
Lunds Universitet

Innehållsförteckning

Introduktion.....	7
Sammanfattning.....	8
Bakgrund: personlig assistans som kostnadsfråga.....	10
Problem och frågeställningar.....	12
Hur har vår undersökning genomförts?.....	16
Disposition.....	18
EEM som uppskattningsmetod av felaktiga utbetalningar.....	19
En historik över EEM och fuskuppskattningar.....	20
Kan expertbedömningar uppskatta felaktiga utbetalningar?.....	31
EEM:s tillämpningsområden.....	32
Vetenskaplig evidens.....	34
Tillämpas EEM i enlighet med vetenskaplig evidens?.....	38
Expertuppskattningar: metodologiska rekommendationer.....	39
Heuristik och snedvridning.....	42
EEM:s olika steg.....	48
Expertgruppen.....	50
Frågeställning.....	53
Svar och resultat.....	54
Presentation av resultat.....	57
Dokumentation.....	60
Summering.....	61
Metodologisk analys av EEM:s tillämpning.....	64
Transparens som överordnat värde.....	67
Expertgruppen.....	69
Hur hanteras risker för snedvridningar?.....	74
Vilken frågeformulering ska experterna besvara?.....	79
Underlag för expertbedömningarna.....	81
Resultat.....	90
Billum-utredningen.....	97
Summering av metodologisk analys.....	103
Slutsatser.....	106
Referenser.....	111

Introduktion

SAMMANFATTNING

De senaste årens debatt kring den personliga assistansen och de politiska ansträngningarna att minska dess kostnader har i stor utsträckning ramats in av uppgifter om att flera miljarder kronor består av felaktiga utbetalningar. I den massmediala debatten om behovet av ökad kontroll figurerar de här siffrorna ofta som för givet tagna sanningar. Uppgifterna om de felaktiga utbetalningarnas storlek kommer från ett antal statliga utredningar och rapporter, som alla använder sig av den så kallade ”expertmetoden”, där personer anställda inom statlig kontrollverksamhet får göra uppskattningar av de felaktiga utbetalningarnas storlek. Expert Elicitation Method (EEM), som är dess formella benämning, beskrivs som en etablerad och vetenskapligt underbyggd metod. Den här rapporten granskar vilket vetenskapligt stöd som finns för att använda EEM för att uppskatta felaktiga utbetalningar inom personlig assistans och rimligheten i siffrorna som man kommer fram till. Utifrån en extensiv litteraturstudie, som inkluderar cirka 1 300 vetenskapliga studier, och en djupanalys av centrala publikationer inom forskningen som använder och/eller diskuterar metoden, drar vi två huvudsakliga slutsatser:

(1) Det saknas evidens för att EEM är en lämplig metod för att bedöma felaktiga utbetalningar inom personlig assistans och andra välfärdssystem. Av de cirka 1 300 publikationer som ingår i vårt material hittar vi inte något exempel på att metoden används för att bedöma mörkertal, välfärdsrelaterade frågor, eller ens för att besvara en samhällsvetenskaplig frågeställning. Det är en metod som är utvecklad för att uppskatta fundamentalt annorlunda typer av osäkerhet, framförallt utifrån naturvetenskaplig fakta, varför det helt saknas erfarenheter och kunskap kring hur metoden kan användas för att uppskatta mörkertal i socialförsäkringssystemen.

(2) De statliga utredningar som använt sig av EEM avviker så kraftigt från de vetenskapliga rekommendationerna om hur metoden bör tillämpas att dess resultat, sedda ur ett vetenskapligt perspektiv, inte ger någon meningsfull kunskap om felaktiga utbetalningar inom personlig assistans.

Vi menar därför att det inte är lämpligt att använda bedömningar som genererats med EEM som underlag för policybeslut, i synnerhet eftersom dessa beslut får stora konsekvenser för personer som är berättigade till personlig assistans. Ur ett demokratiperspektiv menar vi att statliga myndigheter och regering dels har ett ansvar att bidra till en debatt som utgår från vetenskaplig

evidens och dels bör grunda sina åtgärder i vedertagna metoder för hur säker kunskap kan nås.

BAKGRUND: PERSONLIG ASSISTANS SOM KOSTNADSRÅG

Under de senaste åren har de funktionshinderpolitiska besluten kretsat kring hur kostnadsökningarna inom den personliga assistansen kan bromsas. En parlamentarisk utredning har fått i uppdrag att ta fram förslag på besparingar och ytterligare en utredning ska se över fusk i socialförsäkringarna. Regeringen har fryst och sedan sänkt ökningen av schablonbeloppet för varje beviljad assistansstimme och i 2016 års regleringsbrev till Försäkringskassan (Socialdepartementet 2015) ges myndigheten i uppdrag att bryta kostnadsutvecklingen inom assistansen. Detta läggs på en rad förändrade lagtolkningar, från 2006 och framåt, som har resulterat i att personer har förlorat sin assistans och att inflödet av nya assistansanvändare har dämpats (Näsman 2016 s. 6). Sedan Lagen om Stöd och Service för vissa funktionshindrade (LSS) infördes 1994 har utgångspunkten för funktionshinderpolitiska insatser varit att behoven ska styra kostnaderna. Regeringens retorik kring personlig assistans handlar idag istället nästan uteslutande om skenande kostnader. Inom funktionshinderrörelsen

beskrivs läget som krisartat och den rådande politiska linjen som en utarmning av assistansreformens ursprungliga intentioner.

Under sommaren och den tidiga hösten 2016 börjar massmedia uppmärksamma konsekvenserna – hur personer med uppenbara behov av stöd och service får sin assistans indragen och hur detta drabbar assistansanvändare och anhöriga. Samtidigt framhärdar regering och statliga myndigheter i att kostnadsutvecklingen inom assistansen är utom kontroll och måste brytas, att det finns ett omfattande ”överutnyttjande” av assistansersättningen och att privata assistansanordnares vinstintresse driver upp kostnaderna (Begler & Lender, 06/09/16; Regnér, 01/06/16). En inte alltför vågad beskrivning är att det funktionshinderpolitiska fokus – som sedan införandet av LSS 1994 framförallt har handlat om lika rättigheter och inkludering – har förskjutits mot besparingar och bidragsfusk. En central aspekt av hur de politiska besluten har motiverats utgörs av återkommande påståenden om att det finns ett stort överutnyttjande av personlig assistans, det vill säga att den beviljade assistansen inte motsvaras av reella behov. Både nuvarande och föregående regering pratade om att över 10 procent av assistansutbetalningarna kan vara felaktiga, vilket motsvarar 2-3 miljarder kronor. I stor

utsträckning har dessa uppgifter satt tonen för den offentliga debatten.

Givet den roll som föreställningar och antaganden om felaktiga utbetalningar har fått för hur den förda politiken legitimeras menar vi att det är viktigt att vetenskapligt granska rimligheten i dessa siffror. Detta är huvudsyftet med vår rapport: att evidenspröva de uppskattningar som återkommer i den offentliga debatten kring hur mycket pengar som läcker ur assistansen.

PROBLEM OCH FRÅGESTÄLLNINGAR

Innan vi specificerar vårt problemområde och våra frågeställningar, vill vi kort ge en bakgrund till assistansen och den politiska diskussion som förs kring vad den kostar. Personlig assistans är en av tio insatser som specificeras i Lagen om Stöd och Service för vissa funktionshindrade (LSS) och har alltsedan lagens tillkomst varit en central del av svensk funktionshinderpolitik. Insatsens grundtanke är att personer med funktionshinder själva ska få bestämma över det stöd man behöver, bland annat genom att man själv väljer vem som ska utföra assistansen. I kontrast till segregerande insatser, såsom särskilda boenden eller dagcenter, är idén att personer med funktionshinder, med lagtextens formulering,

ska kunna ”leva som andra”. Till exempel innebär detta att personlig assistans ska göra det möjligt för personer med omfattande funktionshinder att kunna ha ett eget boende, arbeta, skaffa familj och leva som fullt integrerade samhällsmedborgare. Utan att fördjupa oss i assistansersättningens teknikaliteter innebär insatsens konstruktion att assistansberättigade behovsprövas och får ersättning från det offentliga, men själva väljer vem som ska utföra assistansen. Oron kring stigande kostnader och felaktiga utbetalningar grundas i uppfattningen att en förhållandevis stor andel av den assistansersättning som betalas ut inte motsvaras av verkliga stödbehov, eftersom människor simulerar eller överdriver behov eller eftersom privata assistansanordnare bidrar till att den enskilde får fler timmar än hen har rätt till.

I den här rapporten har vi fokuserat på den metod som används för att bedöma volymen av felaktiga utbetalningar inom assistansen. Som vi kommer att visa bygger i princip alla de siffror som refereras, av såväl sittande som tidigare regeringar och statliga myndigheter, på den så kallade ”expertmetoden”, eller Expert Elicitation Method (EEM), som är den vetenskapliga benämningen. Metoden har också använts för att uppskatta felaktiga utbetalningar inom andra svenska socialförsäkringssystem. Mot bakgrund av såväl vetenskapliga

som samhällsövergripande debatter om relationen mellan forskning och politiskt beslutsfattande blir det intressant att undersöka hur myndigheter och politiker använder sig av vetenskapliga metoder för att motivera och utforma politik. Som vi kommer att visa finns det i det aktuella fallet ett stort glapp mellan den evidens som finns för EEM och hur metoden faktiskt tillämpas, vilket ytterst reser frågor kring demokratisk legitimitet. En snabb slagning i ett antal stora forskningsdatabaser visar dessutom att EEM är ett vanligt förekommande tillvägagångssätt för att till exempel bedöma risker vid kärnkraftsverksolyckor, byggnationers hållfasthet och för att göra prediktioner kring klimatförändringar. Däremot verkar metoden mer eller mindre aldrig användas för att uppskatta mörkertal inom socialförsäkringen, eller för den delen för att besvara andra utpräglat samhällsvetenskapliga frågeställningar. Detta kommer också att bekräftas av vår systematiska litteraturstudie, som spänner över cirka 1 300 vetenskapliga publikationer. Ändå är detta den metod som just nu ligger till grund för hur politiken inom assistansområdet formas.

Om vårt första problemområde handlar om huruvida EEM kan användas för att bedöma storlek på assistansfusk och överutnyttjande, så handlar det andra om i vilken utsträckning tillämpningen av metoden

överensstämmer med de vetenskapliga rekommendationerna. När myndigheter och politiker försvarar EEM så understryker de ofta dess vetenskapliga legitimitet. En fundamental aspekt av vetenskaplighet är dock att man begagnar sig av de erfarenheter, lärdomar och misstag som andra har gjort tidigare. En sådan kollektiv inläring inom vetenskapssamhället har lett fram till att det idag finns en omfattande metodlitteratur som handlar om hur EEM *bör* tillämpas för att ge så tillförlitliga resultat som möjligt. I vilken utsträckning går dessa lärdomar igen i bedömningarna av de felaktiga utbetalningarna inom assistansen? Använder man sig av de metodologiska rekommendationer som forskarsamhället har kommit fram till? Även här kommer vi att visa på ett stort glapp mellan hur EEM är tänkt att tillämpas och hur metoden rent konkret används för att uppskatta volymen av de felaktiga utbetalningarna.

Med andra ord har vår undersökning vägletts av två huvudsakliga frågeställningar:

1. Är EEM en lämplig metod för att göra uppskattningar av felaktiga utbetalningar inom den personliga assistansen?
2. Tillämpas EEM på ett sätt som ger tillförlitliga resultat?

Detta är alltså inte en undersökning om huruvida det

finns fusk och överutnyttjande inom assistansersättningen eller en granskning av regeringens och myndigheters ställningstaganden och agerande inom området mer generellt. Det är inte heller en undersökning av hur stort överutnyttjande och fusk *egentligen* är. Istället riktar vi vårt fokus mot vad som för tillfället betraktas som *kunskap* om överutnyttjande av personlig assistans och huruvida vi okritiskt bör acceptera denna.

HUR HAR VÅR UNDERSÖKNING GENOMFÖRTS?

Metodologiskt är vår undersökning en vetenskaplig litteraturstudie, som syftar till att sammanställa forskningsläget om EEM och utifrån denna sammanställning analysera metodens tillämpning av svenska myndigheter och utredningar. Vi gör med andra ord inte anspråk på att utvärdera EEM generellt eller att bidra till att metoden utvecklas, utan fokuserar på det givna kunskapsläget satt i relation till den svenska funktionshinderpolitiken.

Våra två frågeställningar förutsätter olika analytiska hållningar. För att besvara den första frågan måste vi få en bred överblick inom vilka områden EEM används och vilka generella rekommendationer som finns kring metodens användningsområden. Detta förutsätter stora kvantiteter av publicerade studier, där bredd i analysen

har getts företräde framför djup. Vår ambition har här varit att ringa in så många studier som möjligt som använder sig av EEM (vi kommer att specificera sökverktyg och söktermer senare i rapporten). Dessa studier har därefter analyserats, i huvudsak genom läsning av abstract och nyckelord, för att avgöra inom vilka områden metoden används.

För att besvara vår andra frågeställning behöver vi först ringa in de centrala metodologiska lärdomarna vid utformandet av EEM-processer. Här har vi arbetat med att identifiera så kallade nyckelreferenser, det vill säga texter som av forskarsamhället anses som viktiga, som återkommande refereras och som andra studier har byggt vidare på. Detta är ett väletablerat sätt att summera forskningsläget inom olika områden och ett standardförfarande i utformandet av vetenskapliga forskningsöversikter. Ett stort antal av de texter vi kommer att diskutera i vår analys utgörs av den här typen av centrala texter om EEM. Dessutom är vår uppfattning att forskningsfältet ger en förhållandevis sammanhållen bild av grunderna för hur EEM bör användas, vilket så klart gör arbetet att ringa in de centrala metodologiska rekommendationerna lättare. Dessa rekommendationer utgör i sin tur utgångspunkten för vår analys av hur metoden använts av svenska myndigheter och utredningar,

där vi kontrollerar i vilken utsträckning tillämpningen av EEM överensstämmer med rådande forskningsläge.

DISPOSITION

I nästa kapitel kommer vi att inleda med att ge en överblick av vilka utredningar och myndigheter som har använt sig av EEM. Därefter presenterar vi en extensiv litteraturöversikt för att besvara vår första frågeställning, som alltså handlar om EEM:s tillämpbarhet. Vi går därefter vidare med vår andra frågeställning, där vi först summerar forskningsläget kring hur EEM bör tillämpas, för att sedan jämföra med hur metoden har tillämpats i praktiken. I vårt slutsatskapitel summerar vi de centrala slutsatserna av vår undersökning.

EEM som uppskattningsmetod av
felaktiga utbetalningar

EN HISTORIK ÖVER EEM OCH FUSKUPPSKATTNINGAR

Under december 1994 ger regeringen i uppdrag åt Riksrevisionsverket (RRV) att kartlägga omfattningen av "fusk med bidrag av social karaktär" (RRV 1995, s. 9). De välfärdssystem som särskilt granskas är förtidspension, a-kassa, bidragsförskott/underhållsbidrag, sjukpenning och tillfällig föräldrapenning. Assistansersättningen faller av naturliga skäl utanför studiens ramar eftersom personlig assistans enligt LSS införts bara ett år tidigare. Utredningen, som är den första mer omfattande kartläggningen av fusk inom svenska välfärdssystem, bedömer att omkring fem till sju miljarder kronor utgörs av felaktiga utbetalningar. (Lundström 2011, s. 27) Metodologin beskrivs av RRV som ett pussel, där olika utredare och externa experter väger samman och analyserar tillgänglig data, indicier och uppskattningar, vilket utmynnar i vad som beskrivs som en grovhuggen bedömning av fuskets totala omfattning. RRV genomför dock inte någon EEM-process, i alla fall inte i begreppets mer formella mening. RRV:s uppskattning av välfärdsfuskets utbredning bygger istället på tre empiriska och analytiska grundpelare: (1) för det första studeras tidigare forskning och nationella studier samt interna analyser gjorda av de berörda myndigheterna och organisationerna. (2)

För det andra analyseras resultat från liknande studier och uppskattningar i andra länder, särskilt de nordiska. (3) Den sista och sammanfogade delen av studien utgörs av uppskattningar gjorda av ekonomen Stefan Fölster på Industrins Utredningsinstitut samt av RRV egna bedömare, som utifrån studiens material och utgångspunkter gör uppskattningar av fuskets omfattning. RRV poängterar dock att ”de iakttagelser som gjorts av oss i detta uppdrag inte heller ger något underlag för väl grundade estimat” (RRV 1995, s. 115-6). Slutsatserna har istället karaktären av att vara överslagsmässiga beräkningar, vilket ”borde vara så långt man kan komma i en sådan här studie” (RRV 1995, s. 116). I kölvattnet av rapporten blossar en debatt om den svenska ”fuskkulturen” upp, men i takt med att svensk ekonomi återhämtar sig från krisåren i början av 1990-talet och att arbetslösheten sjunker avtar också den offentliga debatten om välfärdsfusket. (Lundström 2011, s. 92)

Under ett par år efter publicerandet av RRV:s rapport råder relativ stiltje i debatten om fusk och överutnyttjande av välfärden. Det är först när Kristdemokraternas dåvarande partiledare Alf Svensson under 2002 års valkampanj proklamerar att det troligen förekommer omfattande fusk inom sjukförsäkringen som det blåses nytt liv i rapporteringen. (Johnson 2010, s. 111)

Debatten blir varaktig och förstärks bland annat av att SVT:s Uppdrag granskning under 2005 sänder en serie reportage om välfärdsfusk. Under samma period genomför det Moderata samlingspartiet ansiktslyftet som resulterar i ”Nya Moderaterna” och Alliansen tar form. En central aspekt av den nya borgerliga strategin är att utlova hårdare tag mot välfärdsfuskare och bidragsberoende. Det ökade trycket på reformer och kontroller får den sittande socialdemokratiska regeringen att vidta åtgärder. I vårpropositionen 2005 sätts ett omfattande åtgärdsprogram som syftar till att stävja välfärdsfusket. Extra resurser avsätts bland annat till Försäkringskassan, öronmärkta för anställning av mer personal som bekämpar fusk och överutnyttjande. För att undersöka fuskets orsaker, omfattning och för att ta fram effektiva motåtgärder, tillsätts också den statliga utredningen Delegationen mot felaktiga utbetalningar (FUT-delegationen) (Lundström 2011, s. 115-116; Johnson 2010, s. 222). FUT-delegationens uppdrag skiljer sig från RRV:s tio år tidigare, bland annat i avseendet att fler system (totalt 60 stycken) granskas, men också i att FUT-delegationen gör en distinktion mellan avsiktliga och oavsiktliga fel begångna av bidragstagaren, vilket inte gjordes i RRV:s rapport från 1995. Detta medför att de två rapporternas resultat om det

medvetna fuskets omfattning blir närmast omöjliga att jämföra (Lundström 2011, s. 16).

Det är FUT-delegationen som för första gången uttryckligen använder EEM som metod för att bedöma omfattningen av felaktiga utbetalningar, inklusive mörkertal, inom socialförsäkringssystemen. Metoden används inom sexton ”djupstuderade” trygghetssystem där totalt tio två-dagars expertseminarium genomförs under 2007. Resultaten från de djupstuderade systemen presenteras i en separat rapport i november 2007. Den genomsnittliga omfattningen av felaktiga utbetalningar mellan 2005-2006 beräknas utgöra 6-7 procent av summan av de totala utbetalningarna, vilket motsvarar 15-17 miljarder kronor. Assistansersättningen bedöms vara ett av de system där fusk och fel är mest utbrett: andelen av assistansersättningen som utgörs av felaktiga utbetalningar bedöms uppgå till 10,9 procent, motsvarande 1 241 miljoner kronor. 6,8 procent av de totala utbetalningarna beräknas bero på ”avsiktligt fel av de sökande”, alltså fusk. Detta motsvarar 771 miljoner kronor. Den genomsnittliga bedömningen för samtliga 60 studerade trygghetssystem gör gällande att omkring 4 procent av de totala utbetalningarna är felaktiga och att hälften av dessa beror på avsiktliga fel (FUT 2007:7; Lundström 2011).

Efter regeringskiftet 2006 genomför den nya Alliansregeringen många av de krafttag mot fusk och överutnyttjande som utlovats under valrörelsen. Bland annat införs bidragsbrottslagen och extra resurser tillförs Försäkringskassan för att stärka kontrollverksamheten. Försäkringskassan får också i uppdrag att kontinuerligt genomföra bedömningar av brott och felaktiga utbetalningar inom socialförsäkringens olika förmåner (Lundström 2011, s. 137; Finansdepartementet 2009). Även dessa genomförs med EEM och tycks lita sig mot de metoder och definitioner som tagits fram av FUT-delegationen. Försäkringskassan uppskattar att under 2009 var 3,65 procent (16,4 miljarder) av socialförsäkringens förmånsutbetalningar felaktiga. Sjuk- och aktivitetsersättning, sjukpenning tillsammans med assistansersättningen uppskattas stå för två tredjedelar av de totala felaktiga utbetalningarna, motsvarande ungefär 11 miljarder kronor. Assistansersättningen bedöms vara den ersättning där fusk är mest utbrett; sex procent av ersättningsbeloppen beror på fusk enligt Försäkringskassans experter. Det genomsnittliga fusket för samtliga bedömda ersättningar uppgår till endast 1,3 procent (Socialdepartementet 2010).

Debatten om fusk och felaktiga utbetalningar är fortsatt intensiv under hela 00-talet. 2009 ger regeringen

elva myndigheter i uppdrag att med ”lämplig metod” bedöma omfattningen av felaktiga utbetalningar från de berörda myndigheterna. Samordnande instans för vad som kommer att benämnas ”Samverkansuppdrag mot felaktiga utbetalningar från välfärdssystemen” är Ekonomistyrningsverket (ESV). (ESV 2011:11) I likhet med FUT-delegationen och Försäkringskassans efterföljande uppskattningar, anser ESV att EEM är den bäst lämpade metoden för att bedöma den totala utbredningen av fusk och felaktiga utbetalningar. För att tillåta jämförelse mellan systemen ombeds samtliga delaktiga myndigheter att genomföra expertuppskattningar efter gemensam modell.

Samverkansuppdragets genomsnittliga uppskattning för 2010 är att 3,3 procent av den totala summan har utbetalats på felaktiga grunder. Återigen anges sjukersättningen och assistansersättningen vara de system där felaktigheter och fusk är mest utbredd. 12,2 procent av assistansersättningen bedöms vara felaktigt utbetald, motsvarande 2 147 miljoner kronor. Sju procent av de totala utbetalningarna beror enligt utredarna på avsiktligt fusk. (ESV 2011:11, s. 86-87) Experternas gemensamma osäkerhetsintervall för assistansersättningen är dock stort, mellan 1,8-27,6 procent. Detta motsvarar felaktiga utbetalningar i spannet 309 till 4 859 miljoner

kronor (ESV 2011:11, s. 90). Den stora osäkerheten i bedömningen slår dock inte igenom i medias eller ansvariga politikers framställning av resultaten; flera nyhetsmedier går ut med uppgiften att ESV:s rapport visar att tolv procent av assistansersättningen omfattas av fusk (Hallandsposten, 26/03/11; SVT Nyheter, 24/03/11) och den ansvarige ministern, Maria Larsson (kd), deklarerar att fusk och överutnyttjande inom assistansen utgör ungefär 2 miljarder kronor, alltså cirka 12 procent, men att det kan handla om ännu högre belopp (Dagens Nyheter, 25/03/11; Svenska Dagbladet, 24/03/11; SVT Nyheter, 24/03/11). Att det skiljer drygt 4,5 miljarder kronor mellan det högsta och lägsta uppskattade värdet, vilket är en tydlig indikation på uppskattningens stora osäkerhet, försvinner snabbt ur det politiska och mediala blickfånget.

Framställningen av ESV:s resultat kommer att påverka riktningen för den svenska assistanspolitiken på flera sätt. Regeringens reaktion på ESV:s rapport är att tillsätta en statlig utredning som bland annat ska undersöka hur mycket det fuskas inom den personliga assistansen (SR, 2011).

Eftersom signalerna har varit tydliga, bland annat har ESV tittat på det här och bedömt

att ungefär 12 procent av den totala kostnaden i dag kan handla om fusk och överutnyttjande, då är det angeläget för oss att komma till rätta med det. Fusk ska inte förekomma, säger Maria Larsson. (SR, 2011)

Bara dagar efter ESV:s rapports resultat publicerats tillförordnas justitierådet Susanne Billum till särskild utredare för den statliga utredning som ska bygga vidare på ESV:s rapport. Att ESV:s resultat bidragit till tillsättandet av det som kommer att kallas Billum-utredningen, eller ”Åtgärder mot fusk och felaktigheter med assistansersättningen”, klagörs i kommittédirektivet som föregår utredningen. Som legitimitetsgrund för utredningens nödvändighet hänvisar regeringen till ESV:s resultat:

Av rapporten framgår att den sammanvägda bedömningen av andelen för mycket utbetald assistansersättning som experterna i omfattningsstudierna har gjort, är 12,2 procent av totalt utbetalt belopp, vilket motsvarar 2 147 miljoner kronor år 2010 (Kommittédirektiv 2011:26, s. 9-10).

Även Billum-utredningen fastslår att EEM är den metod

som är bäst lämpad för att uppskatta den totala omfattningen av fusk, felaktiga utbetalningar och överutnyttjande. Dock genomför utredningen inte någon egen EEM-process utan utgår ifrån de resultat som redovisats av ESV (SOU 2012:6, s. 317). I februari 2012 överlämnar Susanne Billum slutbetänkandet till regeringen och ansvarig minister. Utredningens expertgrupp bedömer att fusk och felaktiga utbetalningar uppgår till mellan 9-15 procent av de totala kostnaderna för assistansersättningen 2010, de smalnar alltså av osäkerhetsintervallet från ESV:s rapport. Utöver felaktiga utbetalningar till följd av fel begånget av den sökande eller handläggare bedöms 13-18 procent av den totala kostnaden utgöras av överutnyttjande och övervältring. Totalt gör alltså Billum-utredningens bedömningar gällande att mellan 22-33 procent av utbetalningarna från assistansersättningen utgörs av någon form av felaktigt utnyttjande givet lagstiftningens intentioner eller av övervältring (SOU 2012:6, s. 339).

Efter Billum-rapportens alarmerande höga bedömning ökar kontrollen av assistansanvändarna ytterligare. Den nya regeringen fortsätter på Alliansregeringarnas inslagna väg: under 2015 fortsätter ifrågasättandet av assistansersättningens kostnadsutveckling av bland andra den socialdemokratiska barn-, äldre- och jämställdhetsministern

Åsa Regnér, som genom Försäkringskassans regleringsbrev (Socialdepartementet 2015) ger Försäkringskassan uppgiften att ”bidra till att bryta utvecklingen av antalet timmar inom assistansersättningen” och att ”säkerställa en god kontroll för att motverka överutnyttjande”. Vidare beslutas om att frysa och därefter minska uppräkningen av schablonbeloppen (Näsman 2016, s. 7-8). Genom beslutet att betala ut assistansersättningen i efterskott tas under 2016 ytterligare steg för att säkra att ”ersättningen används på det sätt som lagar och andra föreskrifter anger” (Försäkringskassan 2016). Dessutom meddelar Regnér att en ny utredning ska tillsättas som ska ta fram förslag för hur LSS kan moderniseras. Syftet är att säkra lagstiftningens legitimitet genom att se till så att ”pengarna används till dem de är avsedda för” (Regeringen 2016). I regeringens utredningsdirektiv klargörs att ”kostnaderna för assistansersättningen har ökat kraftigt” och att ”ökningen och ökningstakten är starkt oroande”. I uppdraget ingår även att analysera ”hur insatsen kan begränsas när det gäller antal användare och/eller antal timmar”. (Kommittédirektiv 2016:40, s. 7, 22) Ansvarig för utredningen, som ska presenteras i oktober 2018, är den kristdemokratiske riksdagsledamoten Désirée Pethrus.

Denna historik visar inte bara hur EEM har använts

mot bakgrund av de återkommande diskussionerna om bidragsfusk, överutnyttjande och andra felaktigheter, utan också hur politiska åtgärder och utredningar har motiverats av de resultat som metoden har genererat. Det är mot den här bakgrunden vi nu reser vår första fråga: utgör EEM ett rimligt förfarande för att uppskatta volymen av överutnyttjande och fusk inom assistansen?

Kan expertbedömningar uppskatta
felaktiga utbetalningar?

EEM:S TILLÄMPNINGSSOMRÅDEN

För att få en överblick av kunskapsläget gällande huruvida det går att använda EEM för att bedöma felaktiga utbetalningar har vi genomfört en extensiv litteraturstudie via Lunds universitets sökverktyg LUB-search. Sökverktyget är bland de mest omfattande i Sverige och innefattar cirka 300 000 e-böcker, 200 artikel-databaser och drygt 17 000 e-publikationer. Även om LUB-search inte inkluderar alla vetenskapliga publikationer så täcker sökverktyget stora delar av de senaste hundra årens internationella vetenskapliga produktion och är det primära sökverktyget för forskare på Lunds universitet. Det är osannolikt att den generella bild som framträder vid en sökning i LUB-search skiljer sig från vad andra sökverktyg skulle visa.

EEM är ett systematiserat metodologiskt tillvägagångssätt, varför det också är mycket vanligt förekommande i abstracts och bland de specificerade nyckelorden som i princip alla vetenskapliga publikationer numera preciserar. Genom sökorden "expert elicitation" ELLER "expert judgment elicitation" gavs 1 288 unika träffar. Av dessa utgjordes det stora flertalet av akademiska tidskriftsartiklar. De resterande träffarna bestod av till exempel vetenskapliga böcker, konferensmaterial och

doktorsavhandlingar. Att inkludera ytterligare söktermer, som till exempel ”expert judgment” eller ”expert panel” hade förstås genererat fler träffar, men hade framförallt fångat ett stort antal studier som använder sig av andra metoder som i sammanhanget inte är relevanta.

Vår genomgång av samtliga 1 288 sökresultat visar att EEM är en helt oprövad metod för att besvara samhällsvetenskapliga frågeställningar. Vi hittar inga exempel på att metoden har använts för att bedöma mörkertal, omfattning av överutnyttjande av socialförsäkringssystem, bidragsbrott, bedrägeri eller annan typ av brottslighet. Generellt kan man säga att EEM inte förefaller användas för att göra uppskattningar av mänskligt beteende på aggregerad nivå. Utifrån LUB-search-verktygets egna kategoriseringssystem se vi att vanliga tillämpningsområden exempelvis inkluderar hälsoriskbedömning, klimatförändring, nanoteknologi, ekologi och sjukdomsspridning. För att komplettera den systematiska sökningen i LUB-search har vi även gjort översiktliga sökningar på Google Scholar. I dessa sökningar har vi kombinerat sökorden ”expert elicitation” med bland annat ”mörkertal” ELLER ”hidden statistics” ELLER ”welfare fraud” ELLER ”fraud”, etc. Inte heller i dessa mindre systematiska sökningar har vi hittat studier där metoden används samhällsvetenskapligt, för att bedöma

mörkertal eller icke-rapporterad brottslighet, överutnyttjande, fusk eller bedrägeri.

Det ska återigen poängteras att de sökningar vi gjort inte är uttömmande. Det kan finnas vetenskapliga mörkertalsstudier, eller liknande studier där metoden används, som av en eller annan anledning inte inkluderats i sökresultatet, eller som publicerats i tidskrifter eller av förlag som inte finns med i de databaser som Lund universitet prenumererar på. De sökverktyg som vi har använt inkluderar inte heller offentliga utredningar eller andra typer av rapporter som inte publicerats i akademiska publikationer och som därför inte genomgått akademisk expertgranskning. Vad våra sökningar däremot visar är att EEM inte är en vedertagen samhällsvetenskaplig metod och att det därmed saknas vetenskaplig evidens för att det är en lämplig metod för att göra bedömningar av mörkertal eller felaktiga utbetalningar inom socialförsäkringssystemen.

VETENSKAPLIG EVIDENS

Den begränsade erfarenheten av Expert Elicitation inom samhällsvetenskap och mörkertalsstudier gör det svårt att se vad som utgjort grundval för bedömningen att detta är en lämplig metod för att uppskatta felaktiga

utbetalningar. Vetenskap handlar i stor utsträckning om falsifiering och utveckling genom utvärdering och ackumulerad erfarenhet, där en metod eller teori prövas och kritiseras för att sedan antingen förkastas eller revideras. På så sätt utvecklas kunskap om vad som är gångbart, men också om vilka problem, risker och begränsningar som är förenade med en viss metod eller teori. Avsaknaden av samhällsvetenskapliga studier innebär alltså att är mycket svårt att bygga vidare på tidigare erfarenheter av hur EEM kan tillämpas för att kasta ljus över liknande problemområden.

De svenska rapporter och utredningar som nämnts ovan, i vilka metoden använts för att genomföra omfattningsstudier av fusk och överutnyttjande, refererar inte heller till tillämpningar av EEM inom liknande områden. I FUT-delegationen rapport anges att metoden använts inom "ett flertal antal områden" [sic], bland annat av US Nuclear Regulatory Commission för att bedöma berggrundens lämplighet för förvaring av radioaktivt avfall. Också Statens strålskyddsinstitut anges ha använt metoden för att uppskatta risker avseende radioaktivt avfall (FUT 2007:7, s. 156-157). I de efterföljande omfattningsstudierna genomförda av Försäkringskassan (Socialdepartementet 2010) och ESV (2011) refereras inte heller liknande studier där metoden används, istället

är det FUT-delegationen som man hänvisar tillbaka till. I Billum-utredningen understöds i sin tur metodvalet av att ESV, Försäkringskassan och FUT-delegationen använt metoden¹ (SOU 2012:6, s. 317). Således har ingen av de svenska rapporterna eller utredningarna som använt EEM baserat sitt metodval på tidigare forskning inom det aktuella området. Ur vetenskaplig synpunkt är vår uppfattning att det är behäftat med stor osäkerhet att använda EEM för bedömningar av mörkertal, givet att det egentligen inte finns några tidigare erfarenheter att luta sig mot.

I realiteten innebär tillämpningen av EEM för att uppskatta felaktiga utbetalningar att en metod som utvecklats för att bedöma osäkerhet kopplad till naturvetenskapliga data tillämpas för att besvara en utpräglad samhällsvetenskaplig frågeställning. Detta förfarande strider mot gängse vetenskaplig praxis, där de grundläggande skillnaderna mellan naturvetenskap och studiet av mänskligt beteende är helt centrala för hur samhällsvetenskaplig metodologi har kommit att utvecklas. Med andra ord är det problematiskt att direkt applicera

1. I en intervju publicerad på assistanskoll.se hänvisar en företrädare för Försäkringskassan till att EEM numera används i Norge för att bedöma felaktiga utbetalningar. Detta stämmer förvisso, men även de norska omfattningsstudierna anger FUT-delegationen som inspirationskälla (se Proba, 2013).

metoder som är utvecklade för att besvara naturvetenskapliga frågeställningar för att förstå samhällsproblem. För det första behandlar samhällsvetenskapen mellanmänskliga och sociala fenomen. En grundläggande skillnad mellan att bedöma risknivåer för slutförvaring av radioaktivt avfall och att bedöma omfattningen av fusk och överutnyttjande av assistansersättningen är att den förstnämnda bedömningen inte behöver ta hänsyn till mänskliga faktorer så som sociala strukturer, avsikter eller incitament. Mänskligt beteende följer inga bestämda formler och kan således inte mutas in i generella modeller på samma sätt som till exempel en viss berggrunds hållfasthet och reaktion vid jordskalv kan (Landeta 2005, s. 480).

För det andra, och beslätat med resonemanget ovan, menar vi att det finns fundamentala skillnader mellan de ingångsvärden som används i naturvetenskapliga bedömningar och i de som används för att bedöma felaktiga utbetalningar. I naturvetenskapliga tillämpningar av EEM ligger kända empiriska faktum till grund för expertgruppens utlåtande. Det kan till exempel handla om information om berggrundens beskaffenhet, det radioaktiva avfalllets halveringstid, eller om ett områdes historiska seismologiska aktivitet. Gällande uppskattningarna av mörkertal inom assistansersättningen är

de kända värdena få och övriga värden mycket svårbedömda. Egentligen är den enda empiriska data som experterna har att förhålla sig till de fall av fusk och överutnyttjande som har lett till fällande domar, i övrigt handlar det om indicier och heuristiska antaganden som på centrala sätt skiljer sig från naturvetenskaplig fakta.

Svaret på vår första frågeställning om EEM är en lämplig metod för att bedöma felaktiga utbetalningar, blir i ljuset av ovanstående resonemang ett tydligt ”nej”. Metoden saknar vetenskaplig evidens inom området, den används av allt att döma normalt sätt inte för att besvara den här typen av frågeställningar och den förefaller helt saknas i den vetenskapliga litteraturen om bedömningar av mörkertal och bidragsbrott. Dessutom hänvisar ingen av de utredningar som har använt sig av metoden till evidens för att EEM går att tillämpa inom det här området.

TILLÄMPAS EEM I ENLIGHET MED VETENSKAPLIG EVIDENS?

Vår andra frågeställning handlar om tillämpningen av EEM och huruvida denna följer ur de metodologiska rekommendationer som finns i den vetenskapliga litteraturen. Detta avsnitt är därför indelat i två delar. Först

kommer en sammanfattning av metodlitteraturen där vi preciserar ett antal centrala vetenskapliga rekommendationer för hur EEM bör användas. Därefter analyserar vi tillämpningen av EEM för att uppskatta felaktiga utbetalningar utifrån dessa rekommendationer.

EXPERTUPPSKATTNINGAR:

METODOLOGISKA REKOMMENDATIONER

EEM är en metod för att bedöma risk- eller osäkerhetsnivåer i termer av sannolikhet och kvantitet. Detta görs genom att ”experter” på det berörda området genomför sannolikhetsbedömningar utifrån en strikt och systematisk mall, innan dessa bedömningar vägs samman till ett resultat. Metodtypen har under de senaste decennierna haft en bred tillämpning inom områden såsom biosäkerhet, analyser av hälsorisker samt inom byggnadsteknik och ingenjörsvetenskap (Cooke & Goossens 1999; Landeta 2006; Ouchi 2004, s. 2; Butler et al. 2015, s. 367).

Eftersom EEM är en vedertagen metod inom många områden finns det en stor ackumulerad kunskap kring hur expertbedömningar kan och bör tillämpas för att nå goda resultat. Det ska dock betonas att EEM i strikt mening inte kan förstås som en definitiv metod, utan

snarare som ett samlingsbegrepp för olika, om än besläktade, modeller för systematiskt utvinnande av kunskap från experter. Metoden kan därför användas på olika vis inom olika områden och justeras efter behov, mål och resurser. (Morgan & Henrion 1992, s. 158; se Grigore et al. 2016) Som vi utvecklade ovan är det dock mycket ovanligt att metoden används för att besvara samhällsvetenskapliga frågeställningar och det finns ett stort antal generella rekommendationer som gäller för alla typer av tillämpningar av EEM.

EEM brukar vanligtvis spåras tillbaka till en gren inom det amerikanska flygvapnet, RAND (Research and Development), som efter andra världskrigets slut tog fram metoder för att bedöma scenarion där bevisunderlaget var knapphändigt. Under 40- och 50-talet användes exempelvis den så kallade ”Delphimetoden” för att ta fram riktlinjer för det amerikanska flygvapnets bombräder (Ayyub 2001, s. 3; Knol et al. 2010, s. 2; Morgan 2014, s. 7183) Efter att hemligstämpeln togs bort från RAND-gruppens expertbedömningar började Delphimetoden få stor utbredning under 60- och 70-talet. I sitt standardutförande är alltsedan dess målsättningen med metoden att uppnå konsensus mellan experternas bedömningar i en viss fråga (Knol et al. 2010) Detta åstadkoms genom att experter anonymt och utan kännedom om vilka

andra experter som är involverade, besvarar ett frågeformulär i vilket de uppskattar andelen, antalet eller risken, som frågeställningen berör. (de Franca Doria et al. 2009, s. 811; Rietbergen et al. 2016, s. 169-170) I det klassiska utförandet görs bedömningarna inom ett ”osäkerhetsintervall”, där den i expertens uppfattning största och minsta tänkbara andelen anges. Därefter påbörjas det som vanligtvis benämns som ”Delfi-rundor”, vilket innebär att experterna i projektet bedömer de andras uppskattningar och ger kritik gällande om bedömningarna kan förbättras. Experterna får därefter möjlighet att justera sina bedömningar, varefter en ny runda av utvärdering och kritik följer. Proceduren upprepas vanligtvis till konsensus om osäkerhetsintervallet uppnåtts. I det fall någon eller några experter fortfarande har en avvikande bedömning presenteras dessa med en summerande argumentation och motivering. Noterbart är att hela processen ofta sker anonymt (Ayyub 2001, s. 3-4; de Franca Doria et al. 2009, s. 811).

Just målsättningen att nå konsensus har varit ett återkommande föremål för kritik, menar Morgan:

En Delphi är extremt effektiv när det gäller att nå konsensus, men den här konsensusen är inte baserad på ett genuint samförstånd,

utan är snarare ett resultat av starkt grupptryck mot konformism (2014, s. 7183).

För att motverka effekterna av likriktning är det därför vanligt att konsensuskravet slopas och under de senaste decennierna har det också blivit allt mer vanligt att genomföra öppna EEM-processer, där experternas bedömningar diskuteras och analyseras under rundabordssamtal, lett av en i frågan initierad samtalsledare.

Sammanfattningsvis är alltså EEM en metod som syftar till att kanalisera expertkunskap för att komma fram till estimeringar i frågor där det råder stor osäkerhet, varför andra vedertagna vetenskapliga metoder som bygger på empiriska iakttagelser inte kan användas. För att få en större inblick i vilka insikter som finns att hämta från forskningslitteraturen kommer vi nu att fokusera några huvudsakliga teman för metodologiska rekommendationer. Även om EEM går att använda sig av på olika sätt finns det lärdomar som återkommer och det är dessa vi kommer att ringa in.

HEURISTIK OCH SNEDVRIDNING

Som nämnts är uppskattningar, förutsägelser och sannolikhetsbedömningar fundamentala aspekter av alla typer

av EEM. En sammanfattande benämning för den här typen av resonerande och problemlösning är *heuristik*, ett begrepp med ursprung i grekiskan som har den ungefärliga betydelsen ”hitta” eller ”upptäcka”. Heuristik är samlingsbegreppet för metoder och kognitiva processer där man med bristfälliga bakgrundskunskaper söker uppskatta eller lösa ett problem, utan ambitionen att nå ett fullständigt eller korrekt svar. Vi använder dagligen ostrukturerade heuristiska metoder så som tumregler, kvalificerade gissningar, profilering, och fördomar (Ayyub 2001, s. 20-11; Usher & Strachan 2013, s. 812) Heuristisk är med andra ord ett samlingsbegrepp för de kognitiva processer människor tar till när de ställs inför bedömningar grundade i ovisshet. Eftersom människan inte är utrustad med ett avancerat och objektiva mentalt statistikprogram tenderar dock våra bedömningar att påverkas av snedvridningar, eller ”bias”, vilka resulterar i missvisande subjektiva uppskattningar. (Morgan 2014, s. 7177; EFSA 2014, s. 30) EEM strävar efter att begränsa nivån av osäkerhet och subjektivitet genom att systematisera uppskattningsprocessen och kontrollera för olika typer av heuristiska snedvridningar. I litteraturen om heuristik och EEM finns en stor mängd studier om psykologiska snedvridningsfaktorer som kan påverka våra omdömen när vi gör sannolikhetsuppskattningar.

(se O’Leary et al. 2009, s. 381) Vi ska här begränsa oss till att beskriva fem heuristiska fallgropar som på ett allvarligt sätt kan påverka resultatets giltighet; (1) *tillgänglighet*. (2) *förankring*, (3) *motivationsbias* (4) falska grundantaganden och (5) *övertro*.

Tillgänglighet handlar om att ju mer mentalt tillgänglig en viss typ av händelse är desto mer benägen är experten att tillskriva händelsen hög sannolikhet eller stor utbredning. På så vis överskattas sannolikheten för uppmärksammade händelser och situationer, medan mindre spektakulära och omskrivna händelser tenderar att underskattas. (Morgan 2014, s. 7177; Burgman et al. 2006, s. 14-15). Konsekvensen av detta är att ovanliga händelser ofta tillskrivs överdriven sannolikhet, eftersom ovanliga händelser tenderar att få större publicitet (O’Hagan et al. 2006, kap 3; USEPA 2009:37).

Snedvridning kopplad till *förankring* innebär att experter är benägna att hålla fast vid sin initiala uppskattning, även om ytterligare information innebär att uppskattningen borde justeras mer kraftfullt. I förhållande till EEM innebär detta att revideringen som tillåts efter samråd eller utvärdering av andra experter inte får tillräcklig effekt. (Burgman et al. 2006, s. 15) Ett flertal experiment har visat att förankringsproblematik inte bara påverkar experternas egna initiala uppskattningar, utan

att även initiala värden som är fullständigt arbiträra och slumpmässiga, men som givits till expertgruppen som referenspunkt, har en liknande effekt. Utifrån metodlitteraturen blir det med andra ord intressant att resa frågor kring huruvida experternas förförståelse kring felaktiga utbetalningar och tillgänglig bakgrundsinformation kan bidra till snedvridning i bedömningarna.

Snedvridning kopplad till experternas direkta eller indirekta egenintresse av frågeställningens resultat beskriva som en vanlig risk i EEM-processer. Detta innebär alltså att experternas omdöme, medvetet eller omedvetet, påverkas av dessas intressen, vilket ger upphov till *motivationsbias*, där experten till exempel överskattar ett fenomen för att de vill att de ska vara sant eller för att de har ett intresse av ett specifikt policyutfall. Den här typen av bias är till exempel ett gott skäl att avstå från att välja experter för en EEM-process om hantering av kärnavfall som bor på den tilltänkta plats för slutförvar som ska bedömas. I den amerikanska miljöskyddsmyndighetens riktlinjer för Expert Elicitation-processer konstateras att:

[S]nedvridningar orsakade av motivationsbias är bland de mest svårfångade, men också mest centrala snedvridningarna att

ta hänsyn till för EEM som understödjer policybeslut (USEPA 2009, s. 39).

Övertro, eller *overconfidence* uppstår till följd av att experter överskattar sin bedömningsförmåga, vilket tenderar att leda till alltför snäva osäkerhetsintervall och en ovilja att ompröva sina ståndpunkter. (Mosleh & Apostolakis 1988, s. 66-67; Speirs-Bridge et al. 2010) Studier har dessutom visat att experter, i likhet med lekmän, tenderar att överskatta sannolikheten av små värden, och underskatta sannolikheten för större värden (se Smithson 2014, s. 324; EFSA 2014, s. 31).

Problematik kopplad till bias relaterar egentligen till all typ av heuristik verksamhet, där metoden nödgår uppskattningar eller gissningar grundade i bristfällig kunskap eller otillgänglig information. Det finns dock väldokumenterade metoder för att motverka sådan bias och för att minimera heuristikens fallgropar. Vi kommer att återkomma till huruvida tillämpningarna av EEM för att uppskatta felaktiga utbetalningar använder sig av dessa senare i texten. Det ska dock poängteras direkt att argumenten kring bias inte syftar till att belägga att experter gör dåliga uppskattningar eller att problemen de ställs inför är för svåra för att ens försöka. Poängen är snarare att alla människor tenderar att göra den här

typen av misstag vid heuristiska bedömningar, oavsett om man är expert eller inte (Kahneman & Tversky 1977), vilket innebär att det måste finnas en medvetenhet och strategi kring hur heuristiska snedvridningar kan förebyggas (Butler et al. 2015, s. 374).

En relaterad poäng som återkommer i metodlitteraturen om EEM är att det ur ett givet kunskapsläge, präglat av begränsad information, aldrig kan tillkomma mer kunskap endast genom expertskattningar. Målet med EEM är inte att skapa information, utan att sammanställa tolkningar av befintlig information. (USEPA 2010, s. 56) Som Knol et al (2010, s.7) har påpekat innebär detta att EEM:s framgång är beroende av tillgången till oberoende expertis:

Experter kan inte hitta på kunskap som inte än finns på ett eller annat sätt. För när frågor är behäftade med stor osäkerhet, är kontroversiella, [...] eller är av stor betydelse, kan det finnas otillräcklig expertis tillgänglig för att härleda valida omdömen (Knol et al. 2010, s. 7).

Följaktligen, menar författarna, kan det finnas fall då olika försiktighetsåtgärder för att motverka bias inte kompenserar för den faktiska bristen på kunskap.

EEM:s OLIKA STEG

I metodlitteraturen om EEM finns det idag en relativt stor kunskap kring hur uppskattningar genomförs på bästa sätt och hur processen bör utformas för att uppnå så träffsäkra, pålitliga och transparenta resultat som möjligt. Den här litteraturen berör samtliga delar av elicitations-processen: allt från framtagandet av problemet, formulerandet av frågor till experterna, sammansättning av expertgruppen, sammanvägning av resultat och dokumentation och presentation av processen. (Davis et al. 2006; Knol et al. 2010, s. 2; O’Leary et al. 2009, s. 380-382) Av utrymmesmässiga skäl kan vi inte här utförligt redogöra för samtliga överväganden som aktualiseras när en EEM-process designas, utan kommer nöja oss med en översikt som vi anser gör den generella vetenskapliga diskussionen rättvisa och som lyfter fram de metodologiska rekommendationer som det råder konsensus kring.

I litteraturen görs ofta en distinktion mellan två typer av elicitations-processer: den informella och den formella. Den informella är, som namnet antyder, inte så rigorös som den formella och kräver mindre resurser. Det finns inga vattentäta skott mellan informell- och formell elicitation och de båda tillvägagångssätten kan användas

parallellt (Otway & von Winterfeldt 1992, s. 84). En informell EEM har färre experter, begränsad systematik i genomförandet och en mindre utförlig beskrivning av resultat och process, vilket leder till större osäkerhet och risk för snedvridna bedömningar. Av dessa anledningar menar Slottje et al. (2008) att mer begränsade elicitationer endast bör genomföras då till exempel ny information som förväntas reducera den rådande osäkerheten beräknas komma inom en nära framtid eller då stora delar av det empiriska materialet är samstämmigt. Slottje et al. (2008) menar vidare att en fullständig och utförlig EEM, med en skräddarsydd metodprocess och systematiska urval av experter (vi återkommer till innebörden av detta nedan) är motiverad om följande villkor är uppfyllda:

Vissa empiriska data kan anskaffas på rimligt vis.

Variablernas osäkerhet är stor.

Frågan är känslig/omdebatterad och resultatet kan komma att ha reell inverkan på politiskt agerande och beslut.

Dessa villkor stämmer väl överens med hur frågan om de felaktiga utbetalningarna beskrivs.

I takt med att EEM har utvecklats har också en tämligen samstämmig bild växt fram gällande vad en fullständig EEM-process bör innehålla. Vanligtvis beskrivs

en sjustegs-procedur som inom forskningen ofta framhålls som ett slags minimikrav för vilka byggstenar tillämpningar av metoden bör inkludera (Otway & von Winterfeldt 1992, s. 84-85; Cooke & Goossens 1999; Forester et al. 2004, s. 209; Burgman et al. 2006, s. 21; Knol et al. 2010, s. 2). Proceduren kan ibland brytas ner i ännu fler steg, men ser i normalfallet ut enligt följande:

- 1) Definiering av osäkerheterna.
- 2) Bestämning av och elicitationens räckvidd.
- 3) Val av experter.
- 4) Utformning av Elicitationsprotokollet.
- 5) Förberedelse inför Elicitationsmöten.
- 6) Elicitation av expertkunskaper och omdömen.
- 7) Sammanställning av resultat och rapportering.

Vi kommer inte att utveckla alla steg i den här rapporten, utan nöjer oss med att redogöra för de mest centrala komponenterna.

EXPERTGRUPPEN

Gällande expertgruppens sammansättning finns det ett antal rutiner och processer som rekommenderas mer eller mindre universellt, för alla typer av elicitationer oavsett ambitionsnivå och rigorositet.

En central metodologisk rekommendation är att

expertpanelen bör vara sammansatt så att den ger uttryck för ett brett spektrum av åsikter, synsätt, expertkompetenser, och organisatorisk representation (Cooke & Goossens 1999, s. 29; de Franca Doria et al. 2009, s. 812; Butler et al. 2015, s. 376; Rietbergen 2016, s. 170-171). Slottje et al. (2008, s. 34) menar att en homogen expertgrupp tenderar att producera överflödiga information och Ayyub (2001, s. 6) framhåller att expertgruppens heterogenitet är avgörande för resultatets träffsäkerhet och konstaterar att ”expertpanelens mångfald och helhet är essentiell för elicitations-processens framgång”. Knol et al. (2010, s. 7) argumenterar på ett liknande sätt att en diversifierad expertgrupp är av särskild vikt när den initiala osäkerheten är stor och insatserna höga, till exempel då resultaten kan komma att ha inverkan på policybeslut eller vissa grupper kan komma att påverkas särskilt av utfallet. För sådana frågor rekommenderar författarna att formella procedurer används för att utse expertgruppen och för att skapa en balans som avspeglar det aktuella expertområdets olika synsätt (Knol et al. 2010, s. 7).

Det första steget i urvalsprocessen är att precisera vilken typ av expertis som efterfrågas och vilka olika typer av bakgrunder och kompetenser som behövs för att skapa en heterogen grupp som kan belysa den aktuella

frågans olika infallsvinklar. Vidare påpekas att den vetenskapliga litteraturen inom det aktuella området bör när-granskas för att på så sätt skapa en uppfattning om vilka personer som besitter den efterfrågade expertisen, men också för att utröna vari meningsskiljaktigheter kan föreligga, vilket i sin tur skapar ett underlag för en bred och representativ expertgrupp. Välrenommerade forskare bidrar dessutom till elicitationens trovärdighet (Slottje 2008, s. 20-21; Knol et al. 2010, s. 7).

I litteraturen påtalas även att när elicitationen behandlar konfliktfyllda områden eller områden där den vetenskapliga oenigheten är stor, kan det finnas fog att involvera berörda intresseorganisationer i nomineringen av experter (Slottje et al. 2008, s. 20; se även Kasimir et al. 2003; Kloprogge & van der Sluijs 2006). En generell rekommendation är därför också att urvalet av experter görs genom genomskinliga processer, där noggrannhet och systematik är vägledande. Då EEM utförs på uppdrag av eller inom en myndighet förordas att urvalsprocessen genomsyras av transparens och att strikta procedurer för identifiering och urval av experter används. Ayyub konstaterar i sin guide för EEM att ”formella nominerings- och urvalsprocesser ska lägga grunden för lämpliga kriterier för nominering, utväljande och avsättning av experter” (2001, s. 6). Det finns alltid

en risk – som återkommande påtalas i litteraturen – att EEM-processen kan komma att uppfattas som politiskt styrd, eller på andra sätt partisk, varför transparens och systematik är särskilt viktiga ledord för expertbedömningar (Hora & Jensen 2002, s. 6; USEPA 2009, s. 69).

FRÅGESTÄLLNING

En av de aspekter som inom litteraturen framhålls som avgörande för elicitationens träffsäkerhet och tillförlitlighet är att processen utgår från en tydlig frågeställning. Frågeställningen och de parametrar som ska bedömas bör vara formulerade så att ingen tveksamhet kan uppstå som gör att experterna uppfattar innebörden på olika vis. En tvetydig begreppsanvändning riskerar att leda till osäkra och ojämförbara bedömningar och försvårar själva bedömningsprocessen för varje enskild expert, vilket inverkar negativt på resultatens tillförlitlighet (Morgan & Henrion 1992, s. 159; EFSA 2014, s. 29). Forskningen visar att till och med små omformuleringar av en frågeställning kan innebära stora uppskattningsskillnader: i en studie förändrades till exempel den genomsnittliga kvantitativa uppskattningen av en okänd variabel från fyra till femton procent endast till följd av en subtil omformulering av den fråga som experterna hade att ta ställning till (Knol et al. 2010, s. 9).

SVAR OCH RESULTAT

I litteraturen om EEM finns en relativt utbredd diskussion om hur man på bästa sätt kan utvinna så mycket information som möjligt ur expertgruppen. En avvägning som flitigt problematiserats inom EEM-litteraturen är enligt vilken modell experternas svar bör ges. Vi kommer inte här att redogöra för samtliga alternativ, utan nöjer oss med de två som framstår som de mest använda och omdebatterade: *punktskattning* och *osäkerhetsintervall*.

Punktskattning innebär att experterna uppmanas att uppskatta ett värde av den okända kvantiteten. Problematiken med den här typen av svar är bland annat att en punktskattning inte ger någon information om expertens osäkerhet i frågan, vilket ger resultatet karaktären av en ”kvalificerad gissning”. Eftersom EEM vanligtvis används när osäkerheten kring den okända variabeln är stor är bedömningen av själva osäkerheten i uppskattningen en helt central del av metoden. Många forskare som använder EEM menar därför att det egentligen är bedömningen av experternas osäkerhet som är det centrala vid en elicitation, vilket alltså riskerar att gå förlorat vid punktskattning (Slottje et al. 2008, s. 24).

Ett relaterat problem uppkommer när experternas punktskattningar på något sätt ska aggregeras eller

kombineras. En sådan övning kan resultera i att experternas bedömningar behandlas som mätresultat eller direkta observationer, vilket expertbedömningar inte är. På så vis kan punktskattningar felaktigt tolkas som ordinära statistiska eller vetenskapliga resultat, när det i själva verket rör sig om subjektiva osäkerhetsbedömningar. För elicitationer med högre ambitionsnivå brukar olika varianter av osäkerhetsintervallsbedömningar därför framföras som det rimligaste alternativet för hur det sammanvägda resultatet bör presenteras (Burgman et al. 2006, s. 31; Grigore et al. 2016).

Oavsett vilken typ av statistiska mått och uppskattningsmetoder som används klargör litteraturen att själva osäkerhetsbedömningen, eller uppskattningen av den okända variabeln, bara är en del av det resultat experterna förväntas leverera under en elicitation. Till exempel understryks ofta att den argumentation experterna använder för att understödja sina bedömningar också bör ses som en central del av resultatet (EFSA 2014, s. 35). Detta har återigen att göra med att en EEM inte är utformad för att beskriva hur världen ser ut, utan för att beskriva hur osäkra frågor betraktas av experter inom området. Det är därför rekommenderat att experternas bedömningsgrunder och tankeprocesser nogsamt antecknas och görs tillgängliga för offentlig utvärdering,

tillsammans med de nominella bedömningarna (Morgan & Henrion 1992; Knol et al. 2010, s. 9). Genom att låta experterna motivera och klargöra beslutsgrunden för uppskattningarna tvingas de dels överväga sina bedömningar och dels blir möjligheten att kontrollera för olika typer av snedvridningar betydligt större, vilket stärker resultatets tillförlitlighet och legitimitet. (Knol et al 2010, s. 9) På liknande grunder argumenterar det amerikanska miljöskyddsverket att:

eftersom resultat [från EEM] är baserade på subjektiva omdömen finns en risk att de kan uppfattas som godtyckliga. [...] Men genom att utvärdera den kvalitativa diskussionen som sammanfattar expertens resonemang, kan man bedöma om expertens logik är rimlig och konsistent med tillgänglig empiri och teori (USEPA 2009, s. 25).

I Hora och Jensens metodologiska rapport om EEM som riskbedömningsmetod för slutförvaring av radioaktivt avfall konstateras, i enlighet med resonemanget ovan, att det mest grundläggande kravet man kan ställa på expertomdömesförfaranden är att ”experternas argument dokumenteras och görs tillgängliga för offentlig utvärdering” (2002, s. 1).

Ytterligare ett skäl till detta är ovan nämnda risker för bias. Det är av stor vikt för bedömningens reliabilitet och validitet att experterna, på ett för dem så intuitivt sätt som möjligt, kan ange och motivera sina uppskattningar. (Kynn 2007, s. 251; Morgan 2014, s. 7178) En stor mängd studier har genom åren visat hur experter så väl som lekmän på olika vis gör felaktiga bedömningar kopplade till begränsad förmåga till sannolikhetsbedömning (se Brugman et al. 2006; Kynn 2008). Detta behöver emellertid inte leda till ett kategoriskt avfärdande av metoder som nyttjar sådana bedömningar. Snarare bör begränsningarna uppmärksammas i EEM-processen och bli vägledande för hur metoden används.

PRESENTATION AV RESULTAT

Expertomdömen kan antingen redovisas individuellt eller sammanfogas till en sammanslagen uppskattning. Det finns i nuläget inte någon egentlig konsensus bland forskare under vilka omständigheter och på vilket sätt sammanfogning ska genomföras. Som poängteras i litteraturen kan det finnas ett intresse hos beslutsfattare och uppdragsägare att sammanräkna experternas individuella bedömningar, så att resultaten kan presenteras som ett tydligt beslutsunderlag. Det finns dock risker

och problem kopplade till detta (Ayyub 2001, s. 3; USEPA 2009, s. 29; Martin et al. 2012). Exempelvis är det inte givet att de utvalda experternas bedömningar är representativa för hela spektrumet av potentiella bedömningar (Keith 1996, s. 139). Det snittvärde som skapas då expertuppskattningarna räknas samman riskerar att bli missvisande om det tolkas som en statistisk sammanställning av områdets expertis. Som Cooke och Goossens (1999, s. 7) poängterar ska subjektiva sannolikhetsbedömningar inte förväxlas med statistiska sannolikhetsbedömningar och därför inte heller behandlas på samma sätt. Varje expert inkluderar, förutom sina kunskaper och erfarenheter, också sina fördomar och biaser i sin bedömning, vilket innebär att en sammanslagning av samtliga bedömningar kan förstärka den initiala individuella biasen – i synnerhet om expertgruppen är alltför homogen. Vad gäller resultat producerade genom EEM finns dessutom inga skäl att värdesätta någon enskild experts bedömning högre än någon annans eller att exkludera så kallade ”outliers”, det vill säga uppskattningar som avviker kraftigt från resten av expertgruppen. Som Morgan påpekar är det:

viktigt att komma ihåg att vetenskap inte handlar om majoritetsbeslut. Ibland är det

uteliggaren i minoritet som i slutändan visar sig vara korrekt. Att bortse från detta kan leda till resultat som inte tillgodoser beslutfattarnas målsättningar (2014, s. 7183).

Å andra sidan har det av somliga framhållits att sammanvägning av experternas enskilda bedömningar gör att fler kan vara villiga att delta som experter eftersom de inte blir personligen knutna till en viss uppskattning. I litteraturen är det ofta medelvägen som beskrivs som önskvärd. Denna innebär att experternas bedömningar presenteras individuellt, men att de inte knyts till expertens namn, utan anonymiseras. Cooke och Goossens (1999, s. 30) föreslår på det här sättet att följande procedur efterföljs vad gäller publicering av experternas namn och bedömningar i EEM-studier:

Experternas namn och organisationstillhörighet publiceras i studien.

All information, inklusive experternas namn och bedömningar, tillgängliggörs för professionell utvärdering (peer review), men inte för allmän spridning.

De individuella bedömningarna görs tillgängliga för allmänheten, men bedömningarna knyts inte till experternas namn, utan tillexempel till en siffra (expert 1, 2, 3 osv).

Experternas bedömningsgrund görs tillgänglig för allmänheten.

DOKUMENTATION

En tämligen okontroversiell rekommendation när det kommer till presentation och resultatbeskrivning inom EEM är vikten av systematisk och omfattande dokumentation av hela processen. Utöver att man på så sätt klargör de överväganden vi diskuterat ovan poängteras också vikten av tydliga redogörelser för tolkningen av resultaten, alltså hur de kan förstås och användas samt vilka dess begränsningar är (Ayyub 2001, s. 3; Slottje et al. s. 2008; EFSA 2014, s. 35). Ayyub menar att:

[e]n omfattande dokumentation över processen är avgörande för att säkra resultatens trovärdighet och acceptans. Dokumentationen bör inkludera fullständiga beskrivning av samtliga steg, de initiala resultaten, de reviderade resultaten, samt de aggregerade resultatens spridning och reliabilitet (2001, s. 6).

SUMMERING

Utifrån resonemangen ovan går det att dra slutsatsen att det i de flesta fall är omöjligt att med säkerhet avgöra en EEM-process träffsäkerhet. Eftersom EEM i grunden handlar om att uppskatta sådant som av olika anledningar är okänt ligger denna problematik i metodens natur. Det är mot denna bakgrund särskilt viktigt att reliabiliteten är stark, vilket här åsyftar möjligheten att reproducera och utvärdera processens alla delar i detalj (Burgman 2006, s. 50). Det är därför av yttersta vikt att en EEM-process kännetecknas av god transparens och kvalitetsmässig dokumentation där experternas underliggande resonemang noga redogörs för. På så vis blir det upp till externa bedömare eller uppdragsgivaren att avgöra i vilket utsträckning resultaten kan anses vara pålitliga och kan utgöra grund för politiska beslut (Lauridsen, Andersson & Pulkkinen 2001, s. 9; USEPA 2009, s. 11).

I litteraturen framhålls att resultaten producerade genom EEM inte bör förväxlas med formell statistisk sannolikhetsanalys, eftersom ”ambitionen att genom gissningar generera sannolikhetsdistributioner, uppbackade av väsentlig osäkerhet [...] riskerar att skapa en falsk känsla av precision” (USEPA 2009, s. 7). Detta

har så klart också att göra med att resultaten vid första anblick kan likna verifierade statistiska resultat. Det är därför centralt att påminna om att expertomdömen byggs på bristfälligt eller osäkert empiriskt material, uppskattningar och personliga erfarenheter. Hade säkrare empiriskt material stått till buds hade inte metoden använts från första början (Ayyub 2001, s. 6). Med detta i åtanke framhåller Morgan och Henrion (1992, s. 168) att ”användandet av experters subjektiva omdömen inte är ett substitut för verklig vetenskaplig forskning”. De menar vidare att metodens resultat kan användas som ett understödjande argument för att konventionell forskning på området är nödvändig (se också Ayyub 2001, s. 5).

Vad gäller EEM-resultat som syftar till att underbygga policybeslut framhåller Morgan och Henrion särskilt vikten av uppföljning:

efter att initiala beslut tagits baserade på expertomdömen, bör man följa upp med tillbörlig vetenskap för att försäkra att policybeslut i framtiden är grundade i den fysiska verkligheten (1992 s. 168).

Ett genomgående tema är med andra ord att EEM förutsätter en stor medvetenhet kring metodens begränsningar och inbyggda osäkerhet. Morgan (2014) uttrycker detta:

väl utförd kan Expert Elicitation utgöra ett värdefullt bidrag till väl underbyggt beslutsfattande. Dåligt utfört kan det leda till värdelösa eller, än värre, vilseledande resultat som riskerar att resultera i illa eller felaktigt underbyggda policybeslut (2014, s. 7176).

I detta citat sammanfattar han många av de huvudsakliga poänger som forskare och praktiker är överens om, nämligen att själva sammanförandet av experter i sig inte är tillräckligt för att metodens resultat ska bedömas som sunda och väl underbyggda. Snarare handlar det om att skapa en process som genomsyras av noggranna överväganden, systematik och transparent och fullödlig dokumentation av processens samtliga steg.

Metodologisk analys av EEM:s tillämpning

I det följande kommer vi att analysera hur det svenska utredningsväsendet har använt EEM för att bedöma omfattningen av fusk och överutnyttjande inom assistansersättningen samt hur tillämpningen förhåller sig till den problematik och de riktlinjer för metoden som beskrivs ovan. Vi kommer att studera hur EEM har använts i FUT-delegationens omfattningsstudie *Vad kostar felan?* (FUT 2007, s. 7), som utgör den sjunde delstudien till vad som utmynnar i slutbetänkandet *Rätt och riktigt – Åtgärder mot felaktiga utbetalningar från välfärdssystemen* (SOU 2008:74); ESV:s *Samverkansuppdrag mot felaktiga utbetalningar från välfärdssystemen 2010* (ESV 2011:11); och betänkandet *Åtgärder med fusk och felaktigheter med assistansersättningen* (Billum-utredningen) (SOU 2012:6).

Innan vi inleder den metodologiska analysen bör något sägas om hur rapporterna och utredningarna förhåller sig till varandra. Det finns nämligen en stor dos av spårbindenhet kopplad till dessa, vilken innebär att det i mångt och mycket är samma metodologi och begreppsdefinitioner som används i samtliga studier och att dessa kontinuerligt refererar tillbaka till varandra.

Delegationen mot felaktiga utbetalningar (FUT) var först med att använda EEM för att uppskatta omfattningen av felaktiga utbetalningar i de svenska

välfärdssystemen. Den metodologiska användning som delegationen presenterar utgör på många sätt en mall och referenspunkt för senare metodologiska avvägningar. FUT-delegationens metodologiska redogörelser är betydligt mer utvecklade och uttömmande än senare rapporters, vilket delvis kan förklaras av att efterföljande rapporter till stor del återanvänder delar av FUT-delegationens metodbeskrivning. ESV:s metodologiska process och genomförande baseras i sin tur till stor del på FUT-delegationens metod och begreppsanvändning (ESV 2011:11, s. 18). Billum-utredningens bedömningar är gjorda utifrån ESV:s omfattningsstudier, där man nöjt sig med att göra korrigeringar av tidigare resultat (SOU 2012:6, s. 20). Billum-utredningen kommer därför behandlas separat i analysen nedan. Försäkringskassans omfattningsrapport från 2010, *Felaktiga utbetalningar och brott mot socialförsäkringen* (Socialdepartementet 2010) tycks å andra sidan inte luta sig direkt mot FUT-delegationens användning av EEM, i alla fall inte uttryckligen. Försäkringskassan tycks istället göra någon form av självständig EEM. Förvisso saknar rapporten det mesta av vad som borde inkluderas i en acceptabel metodredovisning, varför det över huvud taget är svårt att göra någon bedömning av metदानvändningen. Detta innebär också att vi i analysen inte kommer att

behandla rapporten. Det finns helt enkelt inte tillräckligt med underlag att analysera.

TRANSPARENS SOM ÖVERORDNAT VÄRDE

Som beskrivits ovan har metodlitteraturen kring EEM fastslagit ett antal ramar till vilka man bör förhålla sig då man använder metoden. För det första understryks gång på gång att metoden är behäftad med stora osäkerheter och inte heller följer de principer och normer som normalt förknippas med vetenskaplig metod. Detta leder till att resultaten inte kan betraktas som statistiskt säkerställda eller vetenskapliga i konventionell mening. Av dessa anledningar är den metodologiska processens transparens av särskild vikt. För att EEM ska kunna betraktas som ett legitimt beslutsunderlag för politiskt känsliga beslut krävs en hög grad av systematik och genomskinlighet samt att samtliga orsaker till bias och andra felkällor nog beaktas och kontrolleras för.

Vid läsning av rapporterna och utredningarna som använder sig av EEM är bristen på transparens i metodanvändningen ett återkommande tema, vilket vi också kommer att återkomma till vid upprepade tillfällen nedan. Självklart är detta delvis ett resultat av att utredningar och rapporter är en annan genre än

vetenskapliga studier, som normalt ställer lägre krav på reproducerbarhet och transparens. Samtidigt behandlar metodlitteraturen hur EEM kan tillämpas som grund för policybeslut; kraven på transparens syftar i dessa sammanhang inte bara till reproducerbarhet, utan också till att uppnå demokratisk legitimitet. Strängt taget visar litteraturen att just transparens är en så central aspekt av EEM att det blir tveksamt om det ens handlar om den här metoden om denna aspekt har allvarliga brister. På samma sätt som det inte går att ta sig till affären för att be om en kubformad boll går det att vända sig till den vetenskapliga litteraturen för att plocka upp en version av EEM som saknar höga krav på transparens.

Vår sammantagna bedömning är att flertalet av de åtgärder som rekommenderas för att uppnå en god genomskinlighet vid genomförande av EEM saknas i tillämpningarna för att bedöma felaktiga utbetalningar. Det handlar till exempel om grundval för sammansättning av expertgrupper, redovisning av metodologiska vägval samt dokumentation av processen som gör det möjligt att bedöma uppskattningarnas rimlighet. Även om de tillämpningar som vi analyserar här inte är vetenskapligt publicerade, så betonar metodlitteraturen de här aspekterna av transparens också då EEM används av myndigheter och i policyprocesser, i synnerhet då resultaten kan

få konsekvenser för specifika grupper. Det står helt klart att tillämpningarna av EEM i detta avseende inte bygger på en konsultation av litteraturen kring vad som är att rekommendera vid expertuppskattningar.

Bristen på metodologisk transparens har fått konsekvenser för vår analys. I flera avseenden är utredningarna så otydliga med vad de gör och hur de motiverar sina val att det har varit svårt att analysera metodtillämpningen. Vi har fått deducera oss fram till rimliga förklaringar, detaljstudera mycket begränsad information och, i vissa avseenden, helt avhålla oss från att bedöma metodtillämpningen, eftersom beskrivningarna av metodanvändning är så bristfälliga att det inte finns något att analysera.

EXPERTGRUPPEN

För att genomföra EEM för bedömning av omfattningen av fusk och felaktiga utbetalningar inom assistansersättningen lät FUT-delegationen Försäkringskassan tillsätta en expertgrupp. De av Försäkringskassan internt utvalda experterna skulle uppfylla en kravlista som inkluderande ”lång erfarenhet av aktuellt utbetalningssystem”, ”överblick (på läns- och/eller riksnivå) av aktuellt utbetalningssystem”, ”god erfarenhet av kontrollfrågor och

felstatistik”, ”personerna ska komplettera varandra vad avser organisatorisk nivå” och ha ”hög förväntad personlig integritet i sina omdömen” (FUT 2007:7, s. 159). Huruvida dessa krav efterlevdes går dock inte att uttala sig om utifrån materialet som redovisas, då den enda information som redogörs för i rapporten är att experterna är fem till antalet och kommer från Försäkringskassan (FUT 2007:7, s. 171).

ESV:s redogörelse för expertgruppens sammansättning överträffar dock föregångaren vad gäller minimalistisk beskrivning av expertgruppens tillsättande och kompetensområden. ESV:s fullständiga beskrivning av expertgruppernas sammansättning lyder:

[o]mkring 115 experter har deltagit i omfattningsstudierna från sex olika myndigheter och organisationer. Det har varit mellan fem och nio experter per förman. Experterna har exempelvis varit handläggare, specialister, chefer, jurister och personer med kunskap om IT-systemen (2011:11, s. 85).

I en bilaga till rapporten redovisas också namnen och organisationstillhörigheten till samtliga experter som del-

tagit, men det går inte att utläsa vilka experter som hör till vilken bedömningsgrupp eller hur många experter som ingår i respektive grupp. Samtliga deltagande experter är också här kopplade till offentliga organisationer.

I förhållande till den procedur och de krav på transparens som tämligen unisont beskrivs i forskningen om EEM måste både FUT-delegationens och ESV:s hantering av expertgruppens tillsättning bedömas som undermåliga. För det första framhåller forskningen att expertgruppens heterogenitet, organisatorisk mångfald inkluderad, är avgörande för elicitations-processens framgång. Särskilt när kunskapsunderlaget är svagt och då resultaten kan komma att ha inverkan på policybeslut anses det vara särskilt viktigt att expertgruppen är diversifierad och representativ för olika utgångspunkter och synsätt (Ayyud 2001, kap. 6; Knol et al. 2007, s. 7). Någon organisatorisk diversifiering i FUT-delegationens expertgrupp finns emellertid inte; samtliga experter kommer från Försäkringskassan. ESV anger att experterna besitter olika typer av expertis, vilket dock inte kan bekräftas eftersom de specifika expertgruppernas sammanställning inte redogörs för. Dock rapporteras att samtliga experter även i ESV:s omfattningsstudie är knutna till statliga myndigheter och organisationer, vilket begränsar gruppens möjliga grad av heterogenitet och oberoende.

Ur ett metodologiskt perspektiv utgör detta en tydlig felkälla som undergräver den metodologiska styrkan. Då experterna tillhör samma organisation omfattas de av samma organisationskultur och övergripande ledningsstruktur. Det är dessutom personer som sannolikt har träffat varandra, som möjligtvis också känner varandra som kollegor, och som därmed är inblandade i beroendeförhållanden till varandra och till organisationens formella hierarkier. Dessutom är det uppenbart att experter som arbetar med assistansersättningen, eller andra socialförsäkringssystem, kan komma att påverkas av vilka resultat man kommer fram till. Om man till exempel konkluderar att det sannolikt är väldigt låga grader av felaktiga utbetalningar så kommer detta att få återverkningar i Försäkringskassans prioriteringar och resursallokering. Detsamma gäller så klart också i det motsatta fallet, om man uppskattar att det är höga nivåer av felaktiga utbetalningar. Vi kommer att återkomma till detta när vi skriver om motivationsbias nedan. Härvid räcker det att konkludera att tillämpningen av EEM för uppskattning av felaktiga utbetalningar genomgående bryter mot centrala metodologiska rekommendationer gällande sammansättning av expertgrupper på ett sätt som är så pass avvikande från gängse förfarande att det kraftigt underminerar resultatens vetenskapliga legitimitet.

Dessutom framhålls vikten av en systematisk, formell och transparent urvalsprocedur genomgående i litteraturen om EEM. Inget av dessa kriterier kan bedömas som uppfyllda av vare sig FUT-delegationen eller ESV:s samverkansuppdrag. Det är förstås möjligt att de interna urvalsprocesserna har varit systematiska och formella. Detta kan dock inte bedömas eftersom en sådan process inte är beskriven i någon av utredningarna. Bristen på transparens gör således att både expertgruppernas tillförordnande och dess kompetensspektrum blir omöjliga att bedöma. Givet resonemangen ovan går det dock att sluta sig till att eventuell systematik i utnämningförfarandet inte har beaktat det centrala kravet på att skapa heterogena expertgrupper.

Som beskrivs ovan används EEM i huvudsak utanför samhälls- och beteendevetenskaperna. I sådana bedömningar är det i första hand ämneskunniga forskare som tillförordnas som experter. Det är också forskare som inom metodlitteraturen antas delta i expertgrupper, vilket bygger på antagandet att forskare står för evidens och oberoende i relation till frågorna som ska uppskattas, till skillnad från experter som själva befinner sig inom de organisationer som har att hantera dessa sakfrågor. Vi menar mot den här bakgrunden att det är märkligt att ingen representant för vetenskapssamhället ingår i vare sig FUT-delegationens eller ESV:s expertgrupper.

HUR HANTERAS RISKER FÖR SNEDVRIDNINGAR?

En viktig anledning till att vikten av heterogenitet i expertgruppen återkommande understryks i forskningen om EEM är metodens utsatthet för heuristiska fallgror, eller bias. En expertgrupp sammansatt av personer med likande professionell organisatorisk tillhörighet, befattning och bakgrund, är mer benägen att ge uttryck för samma typer av bias än en mer heterogen expertgrupp. Om det till exempel skulle vara så att anställda på Försäkringskassan är mer benägna att över- eller underskatta volymen av de felaktiga utbetalningarna så kommer denna effekt alltså att förstärkas till följd av expertgruppens homogenitet. I en analytisk översikt över EEM-rekommendationer påpekar Patricia Fleming (2001, s. 248) att användare av EEM bör ta in externa områdeexperter för att begränsa potentiell bias och intressekonflikter av just det slag som diskuterats ovan:

För att motverka att expertomdömenas trovärdighet sätts på spel av ”institutionell påverkan, ekonomisk eller professionell vinning, eller av främjandet av en social eller politisk agenda” [...] rekommenderas ett balanserat urval av experter (Fleming 2001, s. 248).

I FUT-delegationens rapport *Vad kostar felen?* redogörs för totalt tolv tänkbara orsaker till bias och hur dessa antas påverka experternas bedömningar (FUT 2007:7, s. 162-164)². Som beskrivits ovan innebär *motivationsbias* att expertens omdöme, medvetet eller omedvetet, påverkas av deras intressen, förväntningar eller förhoppningar av hur resultatet bör se ut. FUT-delegationen anger att motivationsbias endast kan påverka expertbedömningarna till att bli för låga, där tolkningen av motivationsproblematiken är att experterna, i egenskap av företrädare för Försäkringskassan, riskerar att göra överoptimistiska bedömningar för att inte själva framstå i dålig dager (FUT 2007:7, s. 163). Vi vill dock framhålla att andra typer av intressen också måste övervägas. Det kan till exempel antas att förväntningarna från politiker och andra makthavare var att fusk och felaktigheter var omfattande. Särskilt som detta redan hade konstaterats i FUT-delegationens tidigare rapporter. Vidare går det inte heller att bortse från den potentiella betydelsen av uppskattningarnas resultat för Försäkringskassans övergripande uppdrag och förutsättningar. De uppskatt-

2. I ESV:s samverkansuppdrag och Billum-rapporten ägnas mycket begränsat utrymme åt diskussioner rörande potentiella orsaker till bias eller hur de kontrolleras för, vilket möjligtvis speglar att dessa lutar sig mot FUT-delegationens tidigare arbete.

ningar om fusk och felaktigheter som FUT-delegationen rapporterade föregick till exempel stora anslag till myndigheten som var öronmärkta för att utöka kontrollverksamheten. Under 2005 och 2006 avsattes 300 miljoner kronor i regeringens budgetar för att stärka Försäkringskassans kontrollarbete, vilket bland annat ledde till att Försäkringskassans 50 kontrollutredare på två år utökades till 274 (Riksrevisionen 2011:20, s. 33). Samtidigt innebär höga uppskattningar en möjlighet till förbättring i påföljande uppskattningar.

Det finns en omfattande ekonomisk och statsvetenskaplig litteratur som går under samlingsnamnet ”public choice” och som visat att offentliganställda har en tendens att agera för att myndigheten där man arbetar ska få ökade anslag och möjlighet att växa. Vi vill understryka att detta inte handlar om att experter koordinerat och medvetet snedvrider bedömningar. De faktorer för snedvridning som vi beskriver är väl belagda i psykologisk forskning och handlar inte nödvändigtvis om avsiktligt agerande, utan om hur intressen och för-förståelse snedvrider våra omdömen (Se tex Otaway & von Winterfeldt 1991, s. 91; Hora & Jensen 2002, s. 6; Burgman 2006, s. 18). Poängen är snarare att utredningens diskussion om motivationsbias, sedd från ett metodologiskt perspektiv, är otillräcklig och aningslös. Vi vill också vara

tydliga med att vi inte argumenterar för att det de facto förekommit motivationsbias i expertgrupperna. Snarare är vårt argument att bedömningen av vad som leder till motivationsbias är komplex och mångbottnad, varför det är problematiskt att FUT-delegationen gör antaganden gällande i vilken riktning en snedvridning pekar och vilka faktorer som kan påverka. Det är här värt att påminna om den amerikanska miljöskyddsmyndighetens riktlinjer för EEM, där det fastslås att just motivationsbias är bland de mest svårpreciserade och subtila, men likväl centrala, snedvridningar man har att ta hänsyn till (USEPA 2009, s. 39).

En stor risk för bias kopplad till fusk och felaktigheter inom assistansersättningen är det som i litteraturen kallas för *tillgänglighets*-bias. "Tillgänglighet" innebär i det här sammanhanget att populariserade och uppmärksammande skeenden tenderar att tillskrivas en överdrivet stor omfattning. Som redan beskrivits hade debatten om fusk och fel i välfärdsutbetalningar pågått intensivt i flera år redan vid tillsättandet av FUT-delegationen. Också i utredningens delrapporter gjordes bedömningen att assistansersättningen är ett högrisk-system för fusk och felaktigheter samt att både allmänhet och handläggare uppfattade fusk som mycket utbrett. Att debatten om fusk påverkade allmänhetens uppfattning om fuskets

utbredning bekräftas också av FUT-delegationens attitydrapport (FUT 2007:7, s. 31). Vi kan med andra ord sluta oss till att det förelåg uppenbara risker för tillgänglighetsbias när FUT-delegationens expertgrupper sattes att uppskatta de felaktiga utbetalningarna. Av samma skäl är det inte orimligt att snedvridning koppad till tillgänglighet också kan ha påverkat ESV:s expertgrupp och hur Billum-utredningen senare valde att korrigera ESV:s resultat – i synnerhet då FUT-delegationens rapport innebar att ytterligare fokus riktades mot fuskets inom assistansersättningen. Det är mot den här bakgrunden anmärkningsvärt att ingen av rapporterna redogör för eller har vidtagit några åtgärder för att minska risken för tillgänglighetsbias.

En relaterad aspekt är att experter som vanligtvis arbetar med fusk och överutnyttjande riskerar att överskatta fenomenets utbredning (FUT 2007:7 s. 164). Till exempel har Lensvelt-Mulders et al. (2006, s. 306), i en studie som inte använder sig av EEM, argumenterat att uppskattningar av välfärdsbedrägerier gjorda av tjänstemän som arbetar med att utreda fusk tenderar att överskatta fuskets utbredning. Bolger och Wright (1994, s. 14) visar på liknande resultat i en studie som konkluderar att läkare ofta överskattar utbredningen av sjukdomar och åkommor de ofta kommer i kontakt med

i sitt arbete. Den här typen av bias är välbelagda inom psykologisk forskning (Tversky & Kahneman 1973) och bör tas i beaktande när EEM används. Även om FUT-delegationen för resonemang om detta förstärker den homogena sammansättningen av deras expertgrupp risken för just den här typen av snedvridning. Detsamma gäller även för ESV:s samverkansuppdrag, där expertgruppen lider av samma typ av homogenitet.

VILKEN FRÅGEFORMULERING SKA EXPERTERNA BESVARA?

FUT-delegationens experter hade till uppgift att besvara den övergripande frågan: ”Hur stora bedöms de felaktiga utbetalningarna vara i kronor räknat och som andel av de totala utbetalningarna från trygghetssystemen” (FUT 2007:7, s. 60). Felaktiga utbetalningar utgjordes i omfattningsstudierna av tre kategorier: (1) avsiktliga fel begångna av den sökande, (2) oavsiktliga fel begångna av den sökande och (3) oavsiktliga fel begångna av myndigheten. Avsiktliga fel gjorda av myndigheten bedömdes däremot inte.

FUT-delegationen anger att det varit problematiskt för experterna att nå samförstånd i hur avgränsningen mellan de olika kategorierna ska förstås, vilket lett till problem i uppskattningen av fördelningen mellan

kategorierna. Fel kan uppstå i kombination med varandra och i vissa fall kan gränsen mellan vad som är att betrakta som avsiktligt och oavsiktligt vara otydlig för experterna (FUT 2007:7 s. 92, 151). Problematiken förstärks av att ersättningsprövningar i många fall baseras på bedömningar gjorda av handläggare och andra professionella i samråd med den sökande, varför det i grunden kan vara problematiskt att kategorisera ansvaret för vissa utbetalningar. Detta i kombination med experternas icke-samstämmiga bedömningar av vilken typ av fel som bör inräknas inom de tre kategorierna leder till tvetydighet, vilket försvårar jämförelse inom expertgruppen, men också mellan de olika välfärdssystemen. Vissa system, assistansersättningen inkluderad, omfattar större grad av bedömning, vilket också torde leda till större grad av definitions- och kategoriseringsproblematik. I ett särskilt yttrande till ESV:s samverkansuppdrag varnar Försäkringskassan för just den här typen av problematik. Bland annat framhålls att de definitioner och kategoriseringar som använts i samverkansuppdraget, som till stor del är identiska med FUT-delegationens, på grund av dess tvetydigheter inte har tolkats lika av de olika instanserna och att Försäkringskassan därför ”anser det nödvändigt att en tydlig och gemensam definition av begreppet felaktiga utbetalningar tas fram” (ESV 2011:11, s. 158-159).

Som vi visade i metodgenomgången ovan är oklarheter gällande de frågor som ska besvaras metodologiskt graverande. Otydligheter och tveksamheter i frågeställningen och kategoridefinitioner kan enligt forskningen om EEM medföra stora risker vad gäller resultatens precision och tillförlitlighet, då experternas bedömningsprocesser försvåras om de inte bygger på en enhetlig förståelse av vad som ska bedömas (Knol et al. 2010, s. 9; Morgan & Henrion 1992, s. 159; Cooke & Goossens 1999, s. 5).

UNDERLAG FÖR EXPERTBEDÖMNINGARNA

Som vi påpekat i metodöversikten förutsätter EEM ett säkert empiriskt underlag som grund för expertuppskattningarna. Mot bakgrund av denna etablerade sanning inom forskningsfältet lämnar tillämpningen av metoden för att uppskatta felaktiga utbetalningar en hel del övrigt att önska, för att uttrycka det försiktigt. Då det finns en tydlig spårbundenhet mellan de olika tillämpningarna av EEM fokuserar vi här framförallt på underlaget i FUT-delegationens expertbedömningar.

FUT-delegationen presenterar resultaten av sina EEM-bedömningar i sin sjunde rapport, *Vad kostar felen?* Det framgår här att underlaget för expertbedömningarna till

stor del består av samma delegations femte delrapport, *Varför blir det fel? Orsaker till felaktiga utbetalningar från Trygghetssystemen*. I denna utreds olika tänkbara orsaker till felaktiga utbetalningar, vem de beror på, huruvida de är avsiktliga eller inte, samt hur vanligt förekommande felet är i förhållande till andra fel. Dock görs inga uppskattningar om de felaktiga utbetalningarnas utbredning, vilket alltså innebär att orsakerna till felaktiga utbetalningar föregår och ligger till grund för bedömningen av dessas volym. Vi citerar nedan hur detta motiveras:

Det kan förefalla ologiskt att inte omfattningen redovisas först och därefter orsakerna. Skälet till den valda ordningen är att det har funnits mycket begränsat med omfattningsstudier att basera en omfattningsbedömning på och än färre studier avseende mörkertalet. Med mörkertal menas här de felaktiga utbetalningar som finns men som normalt inte upptäcks i den ordinarie handläggningen. Genom att först kartlägga alla befintliga orsaker till felaktiga utbetalningar har en grund lagts för att bedöma omfattningen. Strävan har således varit att

redovisa de totala felaktiga utbetalningarna, inklusive det okända mörkertalet (FUT 2007:7, s. 20).

Experterna kommer med andra ord att använda en typologi över tänkbara skäl till felaktiga utbetalningar som utgångspunkt för sina uppskattningar av dessas volym. Den uppmärksamme läsaren noterar att detta i sig är ett skolboksexempel på tillgänglighetsbias, här skapat genom studiens upplägg. I själva verket är så klart orsaker och omfattning av felaktiga utbetalningar logiskt åtskilda, vilket illustreras om vi förflyttar resonemanget till ett annat sammanhang; att det finns tusentals olika sätt att råna en bank säger ingenting om förekomsten av bankrån. Dessutom består orsaksförklaringarna i FUT-delegationens femte rapport i sig inte av empiriska iakttagelser, utan av ett annat expertseminarium med myndighetsföreträdare (det framgår dock inte av rapporterna om det rör sig om samma experter som senare deltagit i EEM-processen), vilket inte tillnärmelsevis motsvarar de vetenskapliga krav på underlag för EEM som preciseras i den metodologiska litteraturen.

Utöver den femte rapportens analys av felorsaker hade experterna i omfattningsstudien att ta hänsyn till resultaten från de attitydundersökningar som presenterades i

delegationens sjätte rapport *Vem fuskar och varför? Om attityder till bidragsfusk* (FUT 2007:6). Rapportens resultat utgörs av kvalitativa gruppintervjuer med handläggare och allmänheten samt av kvantitativa enkätundersökningar där omkring 1 000 personer från allmänheten och 1 000 handläggare besvarade frågor som berörde bland annat hur omfattande de tror att välfärdsfusket är, samt deras egen attityd till fusk. Resultaten från undersökningen visade att både allmänhet och handläggare trodde att fusket var mycket utbrett, samtidigt som acceptansen för fusk var mycket lågt (FUT 2007:7, s. 30-31). Även här finns ett uppenbart logiskt glapp mellan bakgrundsmaterialet och det som faktiskt ska bedömas, enligt analogin ovan; hur vanligt poliser och allmänhet tror att bankrån är säger i sig ingenting om hur vanliga bankrån faktiskt är.

Utöver FUT-delegationens egna rapporter inkluderades ett antal tidigare rapporter och utredningar i bakgrundsmaterialet till omfattningsstudierna, till exempel Fusk-rapporten från 1995 (som inte berör personlig assistans), en från studie 2005 utförd av Institutet för arbetsmarknadspolitisk utvärdering (IFAU) som bedömde överutnyttjandet av en särskild del av föräldrapenningen, samt en studie av Skatteverket som mellan 2005-2006 genomförde en kartläggning av utbredningen av och

samhällskostnaden av svartarbete i kombination med arbetslöshetsersättning. Slutligen hade experterna också att utgå ifrån intern statistik framtagen inom de specifika systemen, till exempel Försäkringskassans statistik över hur många anmälningar om felaktiga utbetalningar som görs samt hur stor del andel av dessa som leder till fällande dom (FUT 2007:7, s. 33-52).

Som tidigare framhållits är EEM:s träffsäkerhet helt beroende av den empiri som ligger till grund för expertbedömningarna. Är underlaget bristfälligt är det också mer sannolikt att resultaten har begränsad träffsäkerhet. Det huvudsakliga underlaget för FUT-delegationens EEM bestod av tidigare uppskattningar och analyser, snarare än empiriska fakta, samt internstatistik som antyder att de felaktiga utbetalningarna kraftigt understiger de resultatet som EEM-processerna kom att generera. Det magra bakgrundsmaterialet framhölls också i samband med presentationen av resultaten:

Som tidigare framgått finns det mycket begränsat med data för de enskilda trygghetsystemen såväl som för helheten vad gäller felaktiga utbetalningar (FUT 2007:7, s. 59).

Vår sammantagna bedömning är att det underlag experterna hade att basera bedömningarna på var synnerligen begränsat. Detta av två huvudsakliga skäl. För det första består två av underlagsrapporterna av delegationens egna studier. Den ena utgörs av en expertgrupps samlade utlåtande om hur och varför felaktiga utbetalningar uppstår. Den andra en attitydundersökning där allmänheten och handläggare också ombeds uppskatta hur utbrett välfärdsfusk är i termer av ”ganska omfattande” och ”mycket omfattande”, samt bedöma hur lätt det är att tillskansa sig mer ersättning än man har rätt till (FUT 2007:7, s. 30). Hur resultatet av dessa två studier korrelerar med reella nivåer av felaktiga utbetalningar går inte att bedöma, men det finns ingen uppenbar logisk koppling mellan dem. Som Hermansson och Johnson (2007) påpekar är möjlighet till fusk inte att förväxla med verkligt fuskande. För det andra fanns det inga tidigare omfattningsstudier av fusk inom assistansen. Underlaget utgjordes således av uppskattningar och bedömningar av andra välfärdssystem, till exempel för tillfällig föräldrapenning, svartarbete och sjukpenning. Återigen är kopplingen mellan underlaget och det som bedömts att bedöma som mycket svag.

Av de underlag som anges är det endast Försäkringskassans statistisk av felaktiga utbetalningar som utgör

ett konkret och empiriskt säkerställt underlag för experternas bedömningar. I de skattningar Försäkringskassan gjort genom systematiska slumpmässiga kontroller av utbetalningar beräknas 0,8-1,7 procent av assistansersättningens utbetalningar bero på felaktigheter (FUT 2007:7, s. 40, 96; Socialdepartementet 2007). Försäkringskassans (Socialdepartementet 2007) skattning, baserad på statistiska metoder och kända felaktiga utbetalningar är alltså anmärkningsvärt mycket lägre än den bedömning som nås genom EEM av FUT-delegationen, som landar i slutsatsen att de felaktiga utbetalningarna utgör mellan 6-19 procent av det totala beloppet.

Denna diskrepans behöver adresseras för att resultaten av EEM ska vara trovärdiga, vilket inte sker.

Ur ett metodologiskt perspektiv är randomiserade och grundliga stickprovskontroller kontroller mer trovärdiga än expertbedömningar. Detta menar bland andra bedrägeriforskarna Gee, Buttom och Brooks (2010) som framhåller att statistiska metoder ger träffsäkrast resultat när det gäller att mäta fusk och felaktiga utbetalningar. De påpekar också att så kallade "gissningsuppskattningar" resulterar i högst osäkra resultat och möjligtvis kan producera trovärdiga bedömningar av experters uppfattningar, vilket dock är något helt annat än bedömningar

av faktiskt fusk (Gee et al 2010; Brooks et al. 2011).³

Det kunskapsunderlag ESV:s expertgrupp för assistansersättning baserade sina bedömningar på är på många sätt mycket likt FUT-delegationens. I Försäkringskassans inlägga till ESV:s rapport, där man alltså kritiserar valet av EEM, framhålls att det:

[f]ör många förmåner finns en stor osäkerhet i resultatet, detta beror inte på metoden i sig utan avspeglar snarare brist på riskanalyser, kontrollinsatser, fallstudier och andra underlag (ESV 2011:11, s. 159).

Exakt vilka riskanalyser, vilka fallstudier och vilken data som faktiskt fanns att tillgå som underlag för experterna preciseras dock inte i ESV:s rapport. De statistiska indikatorer som upptas i rapporten och som kan tänkas ha utgjort underlag för expertgruppens elicitation är: upptäckta felaktiga utbetalningar, misstänkta bidragsbrott, samt antalet impulser rapporterade från antingen myndighetshandläggare eller allmänheten. En impuls

3. Det ska påpekas att bedrägeriforskarna inte direkt åsyftar Expert elicitation-metoden, utan andra typer av mindre systematiska expert-undersökningar där till exempel välrenommerade och erfarna handläggare ombeds uppskatta omfattningen av fusk och fel.

är i sammanhanget en vag misstanke om oegentligheter som rapporteras till berörd myndighet, vilken själv får avgöra huruvida impulsen bör granskas eller polisanmälas. Även i ESV:s resultat finns en stor diskrepans mellan empiriska iakttagelser och expertuppskattningarnas resultat, som inte heller här motiveras eller förklaras.

Sammanfattningsvis noterar vi att valet av EEM för att bedöma omfattningen av felaktiga utbetalningar genomgående har motiverats med bristen på säker kunskap. Utifrån den vetenskapliga litteraturen om metoden framstår dock detta snarare som ett gott skäl för att avstå från att använda den. EEM är inte en metod avsedd att generera kunskap där det helt saknas empiri, utan ett tillvägagångssätt för att göra systematiska uppskattningar utifrån empiri som är svårtolkad. Mot bakgrund av detta går det att sluta sig till att tillämpningarna av EEM för att bedöma felaktiga utbetalningar inte motsvarar de rekommendationer om bakgrundsmaterial som understryks i forskningslitteraturen om EEM. Då metodvalet legitimeras av just bristen på bakgrundsinformation, framstår det som oklart vilken förståelse som funnits kring vad EEM är för typ av metod och vad som krävs för att nå träffsäkra resultat.

RESULTAT

En viktig del av alla EEM-processer är hur expertgruppens bedömningar sammanräknas och redovisas. Av särskild vikt är att behandlingen av bedömningar och sammanvägda resultat avspeglar EEM:s syfte och begränsningar. I litteraturen avråds därför från att redovisa elicitationens resultat i punktskattningar, alltså i ett enda värde. Det framhålls att informationen om expertens osäkerhet vid bedömningen – vilket i metodlitteraturen beskrivs som det centrala resultatet – försvinner vid ett ensidigt fokus på punktskattningen. Expertuppskattningar riskerar också att misstas för mer konventionella statistiska eller vetenskapliga bedömningar när de redovisas som punktskattningar. Det rekommenderas därför att resultaten av EEM presenteras i intervall som anger hur stor osäkerhet experternas bedömningar är behäftade med.

Det är med ovan i åtanke anmärkningsvärt hur de svenska omfattningsbedömningarna redovisats och hur resultaten har överförts till politiker och media. FUT-delegationen redovisar i rapporten *Vad kostar felen?* (2007:7, s.69-70) både det man kallar för ”bedömt värde” – vilket motsvarar en punktskattning – samt ett osäkerhetsintervall som beskrivs som ett ”mycket troligt

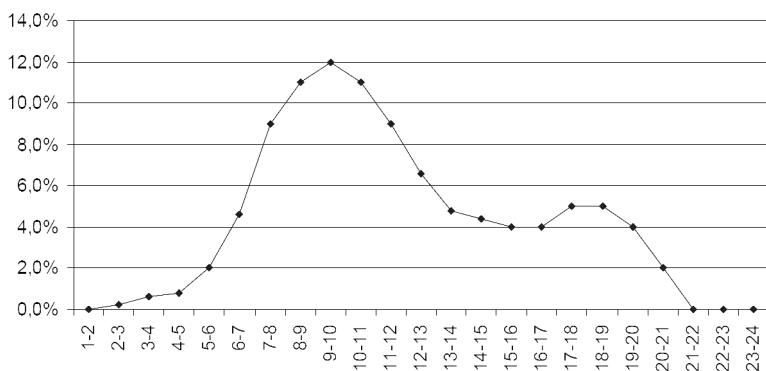
intervall för de felaktiga utbetalningarna”. Experternas sammanvägda bedömning gör gällande att 10,9 procent av assistansersättningen beror på felaktiga utbetalningar och att 6,8 procent av de totala utbetalningarna härrör från fusk. Det ”mycket troliga intervallet” sammanställs till mellan 6 och 19 procent (FUT 2007:7, s. 176). Till ESV:s omfattningsstudier från 2010 har det FUT-delegationen betecknar som ”bedömt värde” bytt namn till ”förväntat värde”, vilket möjligtvis indikerar ännu mer precision (ESV 2011:11, s.13, 80). Det förväntade värdet av fusk och felaktigheter bedöms av ESV:s expertgrupp utgöra 12,2 procent av det totala beloppet och det aggregerade osäkerhetsintervallet är 1,8-27,6 procent (ESV 2011:11, s. 90), vilket ur ett metodologiskt perspektiv är ett exceptionellt brett spann.

Eftersom risken är stor för att de subjektiva sannolikhetsbedömningarna ska uppfattas som precisa när experternas osäkerhetsintervall utelämnas, varnas det kontinuerligt i forskningen om EEM för riskerna med att aggregera experternas enskilda bedömningar i en sammanvägd punktskattning. Ändå var det precis detta som skedde i de politiska och mediala debatter som följde på både FUT-delegationens och ESV:s utredningsresultat. Genom utredningarnas frekventa refererande till dessa tillsynes precisa punktskattningar utelämnades till stor

del osäkerhetsintervallen om 6-19 procent, respektive 1,8-27,6 procent i den offentliga diskussionen. I själva verket ska den genomsnittliga uppskattningen förstås som det enligt experternas aggregerade uppskattningar mest sannolika värdet. Detta är inte att förväxla med ett verkligt värde eller det troliga värdet. För att belysa skillnaderna krävs en utveckling kring hur experternas bedömningsförfarande mer konkret har genomförts, där FUT-delegationens tillvägagångssätt får fungera som exempel.

Inför sina bedömningar utrustades samtliga experter med Excel-ark i vilka de ombads uppskatta sannolikheten för olika intervaller av felaktiga utbetalningar så att den totala summan uppgick till 100 procent (FUT 2007:7, s. 165). Exempelvis hade en expert kunnat uppskatta att det är 5 procents sannolikhet att fusk och felaktigheter uppgår till 4-5 procent av de totala utbetalningarna, 15 procents sannolikhet att andelen är 5-6 procent, 25 procent sannolikhet att fusk och felaktigheter utgör 6-7 procent av det totala beloppet, och så vidare. Det är dessa uppskattningar som utgör experternas individuella osäkerhetsintervall. I diagrammet nedan, hämtat från rapporten *Vad kostar feLEN?* (FUT 2007:7, s. 177), syns experternas aggregerade osäkerhetsintervallsbedömningar. X-axel visar felens andel i procent

av de totala utbetalningarna. Y-axeln visar experternas sammanvägda sannolikhetsbedömning av att felet ligger inom ett visst intervall. Diagrammet visar att det enligt experternas samlade subjektiva bedömningar är omkring 10 till 11 procents sannolikhet att de felaktiga utbetalningarna utgör mellan 10 och 11 procent av de totala utbetalningarna.⁴



Det är uträkningarna som detta diagram är baserat på som gör gällande att 10,9 procent är den enligt experterna mest sannolika omfattningen av felaktiga utbetalningar. Vad som faktiskt visas här är alltså att

4. ”Diagrammet visar den aggregerade sannolikhetsfördelningen med avseende på den totala andelen felaktiga utbetalningar, baserad på de 5 experternas bedömningar. X-axeln visar felets andel i procent av 11 348 miljoner kronor. Y-axeln visar experternas sammanvägda sannolikhet för att felet ligger inom respektive andelsintervall” (FUT 2007:7, s. 177).

10,9 procent är det värde som enligt sammanräkningen av experternas subjektiva bedömningar är mest sannolikt, ungefär 10 procent. Detta motsvarar en högre subjektiv sannolikhetsnivå än för något annat värde. Att göra detta värde till "troligt" eller "förväntat" strider dock radikalt mot EEM-litteraturens rekommendationer. Det detta värde betyder är att fem experter från Försäkringskassan gör bedömningen att det är cirka 10 procent sannolikhet att de totala felaktiga utbetalningarna uppgår till 10,9 procent, vilket så klart innebär att det är betydligt större sannolikhet att någon annan omfattning är den korrekta. Och då förutsätter alltså detta att experterna inte påverkas av bias och i stor utsträckning gör korrekta analyser, vilket vi redan har rest ett flertal frågetecken kring. I ESV:s rapport redovisas inget aggregerat osäkerhetsintervall som korrelerar med sannolikhetsbedömningar, vilket gör det svårt att bedöma hur stor sannolikhet det förväntade värdet 12,2 procent uppgick till.

Ytterligare en problematisk aspekt med de sammanräknade punktskattningarna och osäkerhetsintervallen är att experternas individuella bedömningar, och motiven för dessa, inte redogörs för i någon av utredningarna, vilket som sagt är en återkommande och stark rekommendation i metodlitteraturen om EEM. Det

finns därför ingen möjlighet att utläsa hur experternas uppskattningar skiljdes åt eller hur de interna diskussionerna såg ut. Detta medför också att den aggregerade punktuppskattningen i teorin kan vara ett resultat av att en eller ett fåtal experter har bedömt sannolikheten för värdena kring punktskattningen som mycket högre än övriga experter. Just för att undvika den här typen av oklarheter kring resultaten rekommenderar metodlitteraturen att varje experts bedömningar redovisas tillsammans med deras individuella motiveringar (Morgan & Henrion 1992; Knol et al. 2010, s. 9; USEPA 2009, s.25, 70; Hora & Jensens 2002, s. 1; Cooke & Goossens 1999, s. 30).

Även när det gäller presentation och tolkning av resultat har EEM med andra ord används på ett sätt som kraftigt avviker från de lärdomar som vetenskapliga tillämpningar av metoden har bidragit med. Detta har i sin tur följts av att media och politiker har tolkat punktvärdesskattningarna som statistiskt säkerställda resultat, vilket är en grov feltolkning av vad för typ av metod EEM är och vilken typ av resultat metoden producerar. Som vi nämnde ovan var detta också den tolkning av resultaten som den dåvarande regeringen gjorde, då man använde 12,2 procents fusk och fel som motivering för tillsättandet av Billum-utredningen. I

kommittédirektivet anges följande:

Av rapporten framgår att den sammanvägda bedömningen av andelen för mycket utbetald assistansersättning som experterna i omfattningsstudierna har gjort, är 12,2 procent av totalt utbetalt belopp, vilket motsvarar 2 147 miljoner kronor år 2010 (Kommittédirektiv 2011:26, s. 9-10).

Det är precis den här typen av oreflekterade referenser till punktskattningar som forskningslitteraturen kraftfullt varnar för. Av direktivets formulering framställs ESV:s sammanvägda bedömning som en tämligen precis skattning, med angiven decimal och det totala beloppet preciserat in på miljonen. Det faktum att punktvärdesskattningen i ESV:s rapport benämns som ”förväntat värde” kan på motsvarande sätt inte betraktas som annat än gravt vilseledande. Dessutom måste vi återigen konstatera att bristen på transparens vad gäller metodologisk process och resultatbeskrivning skadar tillitsgraden för både FUT-delegationens och ESV:s bedömningar.

BILLUM-UTREDNINGEN

Vi har ovan analyserat resultaten, underlaget och processen för både FUT-delegationen och ESV:s uppdrag. Vi har dock sagt mycket lite om Billum-utredningens bedömning av omfattningen av fusk inom assistansersättningen. Detta beror i huvudsak på att Billum-utredningen, i motsats till FUT-delegationen och ESV:s samverkansuppdrag, inte använt EEM för att bedöma omfattningen av fusk och felaktigheter. Därför kan Billum-rapportens resultat inte heller analyseras utifrån de metodologiska rekommendationerna på området. Däremot går det att säga en hel del om utredningens förståelse av EEM, givet hur de använder ESV:s resultat.

Vad gäller omfattning av fusk och felaktiga utbetalningar utgick Billum-rapportens skattningar från de bedömningar som expertgruppen inom ESV:s samverkansuppdrag gjort cirka ett år tidigare. Billum-utredningen ansåg dock, mot bakgrund av rapporter och tolkningar för vilka vi redogör nedan, att ESV:s bedömningar behövde justeras (SOU 2012:6, s. 320). Det aggregerade osäkerhetsintersintervallet om cirka 2-28 procent från ESV:s omfattningsstudie ansågs av Billum-utredningen som alltför brett, varför man istället valde att skapa ett nytt intervall löpande från minus 25 procent från

punktskattningen till 25 procent över punktskattningen. Därmed bedömdes de felaktiga utbetalningarna utgöra 9 till 15 procent av de totala kostnaderna. ESV:s punktskattning på 12,2 procent avrundades i Billum-rapporten till 12 procent (SOU 2012:6, s. 320).

Som underlag för dessa korrigeringar hade Billum-utredningen riskanalyser, där incitament och tillvägagångssätt för fusk kartlagts av till exempel Brottsförebyggande rådet och Skatteverket, samt granskningar av handläggnings- och utredningsproblematik. Billum-utredningen genomförde också en enkätundersökning där handläggare ombads uppskatta fuskets omfattning (SOU 2012:6, s. 320). Vidare genomfördes en akt-granskning med målet att samla in och analysera samtliga ärenden som berörde fusk och felaktigheter kopplade till assistansersättningen (SOU 2012:6, s. 231). Utredningens egna beräkningar av fuskets kostnader baserades på 29 ärenden (domar om bidragsbrott/bedrägeri, stämningsansökningar avseende bidragsbrott/bedrägeri och polisanmälningar/pågående utredningar) samt 22 återkrav där ersättning betalats ut felaktigt. Totalt omfattade kartläggningen av fusk och felaktiga utbetalningar alltså 51 fall, för vilka de omstridda beloppen uppgick till 89,2 miljoner, i 2010 års penningvärde (SOU 2012:6, s. 309).

Det är mot bakgrund av dessa fall, tillsammans med enkätundersökningar besvarade av handläggare samt indikationsrapporter och riskanalyser där assistansersättningen återigen beskrivs som särskilt utsatt och svårkontrollerad, som Billum-utredningen gör bedömningen att ESV:s minimivärde nödgar justering från 2 till 9 procent. I monetära termer innebär detta att minimivån för fusk och felaktiga utbetalningar justeras från cirka 473 miljoner kronor, till drygt 2,1 miljarder kronor, räknat med 2010 års totalkostnader. Billum-utredningen gör alltså bedömningen att minimivån för fusk och fel bör justeras uppåt med 450 procent jämfört med ESV:s aggregerade intervall, som vid tidpunkten bara hade ett år på nacken. Detta trots att det fusk som Billum-utredningen totalt sett kunnat konstatera genom akt-granskningen utgjorde 0,4 procent (89,2 miljoner kronor) av 2010 års utbetalningar (SOU 2012:6, s. 255). Det är en häpnadsväckande korrigerings, förutsatt att man fann ESV:s omfattningsstudie så pass tillförlitlig att man kunde behålla punktskattningsvärdet på 12 procent.

Intervalljusteringen av den övre gränsen baserades däremot inte på analys av data och riskrapporter, utan på förmodade kategoriserings- och definitionsproblem hos ESV:s expertgrupp, samt på att ESV:s definition och

kategorisering av felaktiga utbetalningar skiljer sig från den Billum-utredningen brukar:

Det högre värdet i intervallet Försäkringskassan (läs ESV) anger innehåller troligen både fusk och sådant som utredningen bedömer som överutnyttjande och illustrerar svårigheten att dra tydliga gränser mellan vad som är fusk respektive överutnyttjande (SOU 2012:6, s. 320).

Billum-utredningen skiljer nämligen på det man betecknar som ”fusk och andra orsaker till felaktiga utbetalningar” och på ”överutnyttjande” och presenterar två helt separata bedömningar för de två kategorierna (SOU 2012:6, s. 22). ESV gör ingen sådan delad kategorisering utan har som utgångspunkt att den samlade bedömningen ska ”omfatta alla utbetalningar som är fel” (ESV 2011:11, s. 9). ESV nämner över huvud taget inte begreppet ”överutnyttjande” i sin rapport. Eftersom ”överutnyttjande” inte särskiljs från andra typer av fel i ESV:s rapport anser vi att det är mycket problematiskt att bedöma hur stor del av ESV:s aggregerade intervall (2-28 procent) som består av just det som Billum-utredningen kategoriserar som ”överutnyttjande”. Ur ett vetenskapligt perspektiv är det knappast försvarbart att i efterhand definiera och kategorisera resultat som när de framställdes inte baserades på samma definitioner och kategorier. ESV och Billum-utredningen mäter

helt enkelt olika saker vilket borde vara anledning att inte använda den andra utredningens resultat. Detta till trots bedömer alltså Billum-utredningen att ESV:s högre intervall bör justeras från 28 till 15 procent, för att korrigera för utredningarnas olika sätt att mäta felaktiga utbetalningar (SOU 2012:6, 320). Man bör också påminna sig om att ESV:s intervall inte är resultatet av empiriska studier eller statistiska bedömningar, utan av ett litet antal experters subjektiva uppskattningar.

Vidare är de låga och de höga värdena inte, vilket Billum-utredningen tycks anta, resultat av *olika* parametrar eller begreppsdefinitioner, utan visar på experternas osäkerhet i de felaktiga utbetalningarnas utbredning, *givet* vissa parametrar och definitioner. Det är alltså teoretiskt sätt lika troligt att ”överutnyttjande”, på grund av definitions- och kategoriseringsproblem, inkluderas i de låga som de höga värdena i intervallet. Med andra ord borde både minimum-, och maximum-värdena justeras nedåt, för att kompensera inkludering av överutnyttjande. Detta görs inte. Endast maximum-värdet antas vara uppblåst, vilket tyder på att man gör en felaktig tolkning av EEM:s grundläggande logik och bedömningsförfarande. Sammantaget medför detta att det nya justerade intervallet, om plus och minus 25 procent från punktskattningen 12 procent, enligt vår mening kan

betraktas som närapå arbiträrt, givet hur EEM fungerar.

Billum-utredningens tolkning och justering av ESV:s resultat visar på vikten av transparens och tydliga motiveringar av de individuella och aggregerade expertbedömningarna. Hade sådana inkluderats i ESV:s resultatrapport hade måhända Billum-utredningens feltolkning av resultaten kunnat undvikas. Men, eftersom motiveringarna för de aggregerade resultaten, i både FUT-delegationens och ESV:s rapport, var knapphändiga uppkommer risker för missförstånd och feltolkning. Totalt uppgår Billum-utredningens justeringar till värdet om cirka 4,5 miljarder kronor. Att man så kraftigt korrigerar omfattningen av de bedömningar på vilka de i huvudsak bygger sina egna resultat föranleder frågan varför ESV:s resultat ansågs användbara överhuvudtaget.

Slutligen går det också att problematisera Billum-utredningens användning av ESV:s punktskattning, vilken anger 12,2 procent som förväntat värde. Samtidigt som man behåller denna punktskattning väljer man att förändra intervallet. Men en punktskattning framtagen genom EEM är inget fristående värde som på något vis avspeglar ett empiriskt faktum eller en verklig omfattning. Punktskattningen är, som tidigare visat, det värde som ges högst sannolikhet när experternas subjektiva sannolikhetsintervall sammanräknas. Punktskattningen

är alltså ett resultat av intervallet. Ett annat aggregerat sannolikhetsintervall innebär alltid att punktskattningen antar ett annat värde. Att Billum-utredningen ändå utgår ifrån ESV:s punktskattning trots att ESV:s intervall bedöms som totalt felaktiga, tyder på en fundamental brist på förståelse av EEM och vad metodens resultat faktiskt innebär.

Sammanfattningsvis måste vi för det första konstatera att Billum-utredningen inte har någon egentlig metod för att göra en analys av omfattningen av felaktiga utbetalningar. Dessutom tyder de korrigeringar man väljer att göra på att man inte har förstått EEM, vilket leder till slutsatsen att resultaten som presenteras utgörs av mer eller mindre välinformerade spekulationer.

SUMMERING AV METODOLOGISK ANALYS

Vi har i den här analysen av EEM:s användning som verktyg för bedömning av mörkertal och uppskattning av fusk och felaktigheter inom assistansersättningen kunnat konstatera ett antal allvarliga problem.

För det första har vi inte funnit några exempel på att metoden tidigare använts för att göra liknande bedömningar. Det finns enligt vår vetenskap således ingen vetenskaplig erfarenhet att dra lärdom av. Även under

förutsättningen att metoden faktiskt är användbar för att besvara samhälls- och beteendevetenskapliga frågeställningar visar vår analys att det som svenska utredningar och myndigheter faktiskt kallar för EEM inte överensstämmer med den vetenskapliga metoden eller de förfaranden som av forskningen beskrivs som centrala då metoden ska tillämpas av myndigheter. Vi har demonstrerat hur både FUT-delegationens och ESV:s samverkansuppdrags metodologi, på ett flertal punkter, direkt avviker från de rekommendationer som finns i forskningslitteraturen. Till exempel handlar detta om:

Expertgruppernas sammansättning.

Underlaget för experternas bedömningar.

Transparensen i processen.

Oförmågan att konfrontera risker för bias.

Presentation och tolkning av resultat.

Vad gäller de resultat som tillämpningen av EEM har genererat menar vi att de är att bedöma som mycket osäkra – ur ett vetenskapligt perspektiv går de knappast att betrakta som meningsfull kunskap. Vår samlade analys visar att FUT-delegationens och ESV:s användning av EEM är bristfällig till den grad att det måste betraktas som vilseledande att beskriva detta som en forskningsbaserad metod. Gällande Billum-utredningen spelar feltolkning och felanvändning av EEM en stor roll

för de resultat som presenteras, vilket reser frågetecken kring om utredningens förfarande har någon metodologisk grund för sina resultat överhuvudtaget.

Sammanfattningsvis kan alltså sägas att det finns ett stort glapp mellan den säkerhet med vilken de här utredningarnas resultat refereras och den betoning på osäkerhet, transparens och rigorositet som forskningslitteraturen kring EEM kontinuerligt understryker. Man är så klart fri att tro på resultaten ändå, men metodtillämpningen som genererat dessa resultat är knappast evidensbaserad. Snarare misslyckas samtliga tillämpningar som vi har studerat med att tillfredsställa grundläggande metodologiska rekommendationer, vilket om något borde öka osäkerheten i de resultat som presenteras.

Slutsatser

Vår genomgång av den vetenskapliga litteraturen om EEM och vår analys av hur metoden faktiskt används av svenska myndigheter visar på en rad tankeväckande saker.

Som svar på vår första frågeställning menar vi att det saknas vetenskaplig evidens för att EEM kan användas för att förstå det aktuella problemet – det vill säga omfattningen av felaktiga utbetalningar. Detta innebär också att det helt saknas metodologiska rekommendationer kring hur metoden ska användas för att besvara samhällsvetenskapliga frågeställningar. Det övergripande intrycket blir att de olika utredningar och myndigheter som har använt metoden famlar i ett metodologiskt mörker, där de improviserar snarare än konsulterar ”best practice” och auktoritativa källor.

Som svar på vår andra frågeställning har vi visat att den svenska tillämpningen av EEM för att bedöma felaktiga utbetalningar bryter mot flera tydliga metodologiska rekommendationer kring hur metoden bör genomföras. Detta handlar till exempel om vilket experter som deltar, hur dessa väljs ut, vilken typ av frågor som ställs, vilka överordnade värden som ska genomsyra processen samt hur resultaten tolkas och presenteras. I vissa avseenden skiljer sig beskrivningarna av det svenska förfarandet så pass mycket från de rådande rekommendationerna att

det inte är givet att dessa tillämpningar kan klassas som EEM.

En övergripande slutsats är med andra ord att den nuvarande tillämpningen av EEM inte är evidensbaserad; att det saknas skäl att tro att metoden alls kan bedöma storleken på de felaktiga utbetalningarna samt att metodens tillämpning kraftigt avviker från de metodologiska rekommendationerna. Detta bör också få konsekvenser för hur vi värderar den kunskap som genererats. För det första, ur ett vetenskapligt perspektiv bör det understrykas att de uppskattningar som tagit fram med hjälp av EEM inte kvalificeras som tillförlitlig kunskap alls, då man uppenbarligen inte har tagit hänsyn till grundläggande rekommendationer kring hur metoden ska tillämpas. För det andra måste det påpekas att bristerna i förfarande får konsekvenser för säkerheten i bedömningarna. Den metodologiska litteraturen betonar genomgående att EEM redan från början inte ger definitiva och säkra resultat. Denna osäkerhet blir större om vi beaktar tillämpningens metodologiska brister.

Ett särskilt intressant spår i litteraturen har att göra med hur precision, transparens och rigorositet beskrivs som extra viktigt då EEM används i processer som har policyimplikationer, när vissa grupper riskerar att påverkas av bedömningarna och dessa kommer att ligga

till grund för politiskt beslutsfattande (Hora & Jensen 2002, s. 6; USEPA 2009, s. 69, 71; Morgan 2014, s. 7179-7180). Inom assistansområdet har tendensen snarast varit den motsatta: trots att uppskattningarna explicit och återkommande används som underlag för politiska beslut är EEM:s tillämpning av svenska myndigheter långt mycket mindre rigorös, transparent och metodologiskt informerad än vad som rekommenderas i den vetenskapliga litteraturen. Mot bakgrund av den svenska debatten kring personlig assistans, där ”runt 10 procent eller mer” återkommit som ett slags precis minimimått på de felaktiga utbetalningarna, följs läsningen av den vetenskapliga litteraturen om EEM av tilltagande överklighetskänslor. Där litteraturen varnar för säkra slutsatser och betonar osäkerheten, framställs metoden i den svenska kontexten som en tillförlitlig och förhållandevis precis indikator på ett verkligt empiriskt fenomen. Medan litteraturen understryker vikten av legitimitet och transparens, utelämnar de svenska utredningarna ofta grundläggande information om vad det är man faktiskt har gjort. I tillägg till att metodens tillämpning inte kan sägas vara vetenskapligt förankrad, måste vi därför också tillägga att beskrivningarna av resultaten i den offentliga debatten saknar rimlighet. Givet EEM:s inneboende brister samt bristerna i hur metoden

tillämpas säger resultaten oss väldigt lite om volymen av felaktiga utbetalningar inom assistansersättningen.

Referenser

Assistanskoll.se, 2017. "Thomas Falk, talesperson på Försäkringskassans kontrollenhet – "Expertmetoden ger kvalificerade bedömningar av fusket". Tillgänglig: <https://assistanskoll.se/20170317-Falk-Forsekringskassans-kontrollenhet-Expertmetoden-fusket.html>. Hämtad 2017-03-31

Ayyub, B.M., 2001. *Elicitation of expert opinions for uncertainty and risks*. CRC press.

Begler, A. & Lender, A. 06/09/16 "Bolag och brukare har fel om kris i assistansen". *Dagens Nyheter*. Tillgänglig: www.dn.se/debatt/bolag-och-brukare-har-fel-om-kris-i-assistansen/. Hämtad 2017-04-11

Bolger, F. & Wright, G., 1994. "Assessing the quality of expert judgment: Issues and analysis". *Decision support systems*, 11(1), s. 1-24.

Brooks, G., Button, M. & Gee, J., 2012. "The scale of health-care fraud: A global evaluation". *Security Journal*, 25(1), s. 76–87.

Burgman, M. et al., 2006. "ACERA Project 0611 Eliciting Expert Judgements". *Risk Analysis*, s. 1-71.

Butler, A.J., Thomas, M.K. & Pintar, K.D.M., 2015. "Systematic review of expert elicitation methods as a tool for source attribution of enteric illness". *Foodborne Pathogens and Disease*, 12(5), s. 367-382.

Cooke, R.M. & Goossens, L.J.H., 1999. "Procedures

guide for structured expert judgment”. *Project Report to the European Commission*.

Davis, A. et al., 2006. ”Effectiveness of requirements elicitation techniques: Empirical results derived from a systematic review”. *14th IEEE International Requirements Engineering Conference (RE’06)*. IEEE, s. 179-188.

de Franca Doria, M. et al., 2009. ”Using expert elicitation to define successful adaptation to climate change”. *Environmental Science & Policy*, 12(7), s. 810-819.

Ekonomistyrningsverket 2011:11. Gemensam skrivelse. *Samverkansuppdrag mot felaktiga utbetalningar från välfärdssystemen 2010*.

European Food Safety Authority (EFSA), 2014. ”Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment”. *EFSA Journal*, 12(6).

Finansdepartementet, 2009. Regeringsbeslut. *Samverkansuppdrag om utveckling av metoder för och redovisning av resultat av arbetet mot felaktiga utbetalningar från välfärdssystemen*.

Fleming, P., 2001. ”Examining Recent Expert Elicitation Judgement Guidelines: Value Assumptions and the Prospects for Rationality”. *RISK: Health*,

Safety & Environment, 12(1).

Folcker Aschan, Annika, 2011. ”Bedrägerier med assistenter utreds”. *Dagens Nyheter*. Tillgänglig: <http://www.dn.se/arkiv/nyheter/bedragerier-med-assistent-utreds/>. Hämtad 2017-03-31

Forester, J. et al., 2004. ”Expert elicitation approach for performing ATHEANA quantification”. *Reliability Engineering & System Safety*, 83(2), s. 207-220.

FUT-delegationen, 2007:5. *Varför blir det fel? – Orsaker till felaktiga utbetalningar från trygghetssystemen – Resultat från expertseminarier.*

FUT-delegationen, 2007:6. *Vem fuskar och varför? – om attityder till bidragsfusk i Sverige.*

FUT-delegationen, 2007:7. *Vad kostar felet? – Omfattning av felaktiga utbetalningar från trygghetssystemen.*

Försäkringskassan 2016. *Nyhetsbrev*. Tillgänglig: <https://www.forsakringskassan.se/wps/wcm/connect/b69d3e58-7722-44e9-b92f-dd8dcb613892/aldre-nyheter-fran-fkse-om-efterskotts-betalning-161017.pdf?MOD=AJPERES>. Hämtad 2017-04-11

Gee, J., Button, M. & Brooks, G., 2010. *The financial cost of UK Public Sector Fraud. A less painful way to reduce public expenditure*. Center for Counter Fraud Studies, University of Portsmouth.

- Grigore, B. et al., 2016. "A comparison of two methods for expert elicitation in health technology assessments". *BMC Medical Research Methodology*, 16(1).
- Hallandsposten, 2011. "Trygghetssystemet är till för behövande inte fuskare". Tillgänglig: <http://www.hallandsposten.se/%C3%A5sikter/ledare/trygghetssystemet-%C3%A4r-till-f%C3%B6r-beh%C3%B6vande-inte-fuskare-1.1630179>. Hämtad 2017-03-31
- Hora, s. & Jensen, M., 2002. *Expert judgment elicitation*. Swedish Radiation Protection Authority.
- Johnson, B. & Hermansson, A., 2007. "Överutnyttjandediskursen. En innehålls-och konsekvensanalys av diskursen kring överutnyttjande av sjukförsäkringen". *Arbetsliv i omvandling*. Arbetslivsinstitutet.
- Johnson, B., 2010. *Kampen om sjukfrånvaron*, Arkiv förlag.
- Kadane, J. & Wolfson, L.J., 1998. "Experiences in elicitation". *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1), s. 3-19.
- Kahneman, D. & Tversky, A., 1977. *Intuitive prediction: Biases and corrective procedures*. DTIC Document.
- Kasimir, B., Jaeger, C.C. & Jäger, J., 2003. "Public

Participation in Sustainability Science”. *Public Participation in Sustainability Science. A Handbook*, s. 3-36.

Keith, D.W., 1996. ”When Is It Appropriate To Combine Expert Judgments?”. *Climatic Change*, 33, s. 139-143.

Klopprogge, P. & Sluijs, J.P. Van Der, 2006. ”The inclusion of stakeholder knowledge and perspectives in integrated assessment of climate change”. *Climatic Change*, 75(3), s. 359-389.

Knol, A., Sluijs, J.P. Van Der & Slottje, P., 2008. ”Expert Elicitation: Methodological suggestions for its use in environmental health impact assessments”. *RIVM Letter report*.

Knol, A.B. et al., 2010. ”The use of expert elicitation in environmental health impact assessment: a seven step procedure”. *Environmental Health*, 9(1).

Kommittédirektiv 2011:26. *Utredning om åtgärder mot fusk, oegentligheter och överutnyttjande av den statliga assistansersättningen*.

Kommittédirektiv 2016:40. *Översyn av insatser enligt LSS och assistansersättningen*.

Kynn, M., 2008. ”The ’heuristics and biases’ bias in expert elicitation”. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(1),

s. 239-264.

Lag (1993:387) om stöd och service till vissa funktionshindrade, LSS.

Landeta, J., 2006. "Current validity of the Delphi method in social sciences". *Technological Forecasting and Social Change*, 73(5), s. 467-482.

Lauridsen, K., Anderson, K. & Pulkkinen, U., 2001. *NKS/SOS-1 Seminar on Safety analysis*. Nordisk Kernesikkerhedsforskning.

Lensvelt-Mulders, G.J.L.M. et al., 2006. "A validation of a computer-assisted randomized response survey to estimate the prevalence of fraud in social security". *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 169(2), s. 305-318.

Lundström, R., 2011. *Den kalkylerande medborgaren: bidragsfusk i svensk välfärdsdebatt 1990-2010*. Umeå universitet.

Lundström, R., 2013. "Framing fraud: Discourse on benefit cheating in Sweden and the UK". *European Journal of Communication*, 28(6), s. 630-645.

Martin, T.G. et al., 2012. "Eliciting Expert Knowledge in Conservation Science". *Conservation Biology*, 26(1), s. 29-38.

Morgan, M.G., 2014. "Use (and abuse) of expert elicitation in support of decision making for public

policy". *PNAS*, 111(20), s. 7176-7184.

Morgan, M.G., Henrion, M. & Small, M., 1992. *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge university press.

Mosleh, A., Bier, V.M. & Apostolakis, G., 1988. "A critique of current practice for the use of expert opinions in probabilistic risk assessment". *Reliability Engineering & System Safety*, 20(1), s. 63-85.

National Audit Office, 2006. *International benchmark of fraud and error in social security systems*.

Näsman, M., 2016. *Konsekvenser av indragen assistansersättning. En studie om medikalisering och hur kostnader skiftar från samhället till brukaren och dess anhöriga*. Rapport för STIL om förändrat stöd enligt LSS och SFB. Umeå universitet.

O'Hagan, A. et al., 2006. *Uncertain judgments: eliciting experts' probabilities*. John Wiley & Sons.

O'Leary, R.A. et al., "Comparison of three expert elicitation methods for logistic regression on predicting the presence of the threatened brush-tailed rock-wallaby *Petrogale penicillata*". *ENVIRONMETRICS*, 20(4), s. 379-398.

Otway, H. & von Winterfeldt, D., 1992. "Expert Judgment in Risk Analysis and Management: Process,

Context, and Pitfalls”. *Risk Analysis*, 12(1), s. 83-93.

Ouchi, F., 2004. *A Literature Review on the Use of Expert Opinion in Probabilistic Risk Analysis*. World Bank Policy Research Working Paper 3201.

Proba, 2013. *Trygdesvindel i Norge – En kartlegging av fem stønadsordninge*. Arbeidsdepartementet.

Regeringen, 2016. ”Pengar till assistansen ska gå till det de är avsedda för”. Tillgänglig: <http://www.regeringen.se/artiklar/2016/03/pengar-till-assistens-ska-ga-till-det-de-ar-avsedda-for/>. Hämtad 2017-04-11

Regnér, Å., 2016. ”Kostnadskontroll och utveckling går hand i hand”. *Dagens Samhälle*, från www.regeringen.se

Rietbergen, C. et al., 2016. ”Expert elicitation of study weights for Bayesian analysis and meta-analysis”. *Journal of Mixed Methods Research*, 10(2), s. 168-181.

Riksrevisionen 2011:20. *Vad blev det av de misstänkta bidragsbrotten?*

Riksrevisionen 2016:11. *Felaktiga utbetalningar inom socialförsäkringen – Försäkringskassans kontrollverksamhet*.

Riksrevisionsverket (RRV) 1995:32. *Fusk – Systembrister och fusk i välfärdssystemen*.

Slottje, P., van der Sluijs, J.P. & Knol, A.B., 2008. ”Expert Elicitation: Methodological suggestions for

its use in environmental health impact assessments”.
RIVM Letter report.

Smithson, M., 2014. ”Elicitation”. *Introduction to Imprecise Probabilities*, s. 318-328.

Socialdepartementet, 2007. *Felaktiga utbetalningar och brott mot socialförsäkringen – omfattning, risker och information.*

Socialdepartementet, 2010. *Felaktiga utbetalningar och brott mot socialförsäkringen.*

Socialdepartementet, 2015. *Regleringsbrev för budgetåret 2016 avseende Försäkringskassan.*

SOU 2008:74. *Rätt och riktigt. Åtgärder mot felaktiga utbetalningar från välfärdssystemen.*

SOU 2012:6. *Åtgärder mot fusk och felaktigheter med assistansersättning.*

Speirs-Bridge, A. et al., 2010. ”Reducing Overconfidence in the Interval Judgments of Experts”. *Risk Analysis: An International Journal*, 30(3), s. 512-523.

Svenska Dagbladet, 2011. ”Larsson: Handikappfusk för miljarder”. Tillgänglig: <https://www.svd.se/larsson-handikappfusk-for-miljarder>. Hämtad 2017-03-31

Sveriges Radio, 2011. ”Regeringen utreder fusk i assistansersättningen”. Tillgänglig: <http://sverigesradio.se/sida/artikel.aspx?programid=83&artikel=4417862>.

Hämtad 2017-03-31

SVT Nyheter, 2011. "Larsson: Handikappfusk för miljarder". Tillgänglig: <http://www.svt.se/nyheter/inrikes/larsson-handikappfusk-for-miljarder>. Hämtad 2017-03-31

Tversky, A. & Kahneman, D., 1973. "Availability: A heuristic for judging frequency and probability". *Cognitive Psychology*, 5(2), s. 207-232.

U.S. Environmental Protection Agency (USEPA), 2009. *Expert Elicitation Task Force White Paper*. Environmental Protection Agency, Washington DC.

Usher, W. & Strachan, N., 2013. *An expert elicitation of climate, energy and economic uncertainties*. *Energy Policy*, 61, s. 811-821.

Niklas Altermark är fil.dr och forskare i statsvetenskap vid Lunds Universitet.

Han har publicerat en mängd artiklar om svensk funktionshinderpolitik i tidskrifter som "Disability & Society" och "Review of Disability Studies".

Vintern 2017 publiceras hans bok "Intellectual Disability and Citizenship Politics" av Routledge.

Hampus Nilsson är forskningsassistent vid Statsvetenskapliga institutionen, Lunds Universitet.

Han blir färdig med sin Masters-examen med inriktning mot välfärdspolitik våren 2018.

STIL

Personlig assistans och politisk påverkan

