Hur bör forskningsprestationer mätas?

MAGNUS HENREKSON, KRISTIN MAGNUSSON OCH DANIEL WALDENSTRÖM

Magnus Henrekson är chef för Institutet för Näringslivsforskning och professor i nationalekonomi på Handelshögskolan i Stockholm. Magnus.Henrekson@naringslivsforskning.se

Pol mag *Kristin Magnusson* är doktorand i nationalekonomi på Handelshögskolan i Stockholm och forskar kring finanspolitik och optimala skattesystem. Kristin.Magnusson@hhs.se

Fil dr *Daniel Waldenström* forskar vid Institutet för Näringslivsforskning om finansiell utveckling och inkomst- och förmögenhetsfördelning. Daniel.Waldenstrom@naringslivsforskning.se

Vi tackar Niclas Berggren, Anders Björklund, Robin Douhan, Assar Lindbeck, Bo Sandelin och seminariedeltagare vid Handelshögskolan i Stockholm, Industriens Utredningsinstitut och Lunds universitet. Finansiellt stöd har erhållits från Jacob Wallenbergs Fond.

I takt med hårdnande konkurrens om forskningsmedel har behovet av att mäta och utvärdera forskningsprestationer ökat, liksom vikten av att de mått som används verkligen fångar det vi vill mäta. I denna artikel studeras och jämförs utfallet för fem olika mått på en avidentifierad population bestående av samtliga svenska professorer i nationalekonomi födda 1940 eller senare. Undersökningen visar att det mått (KMS) som användes i Tobias Lindqvists (2003) uppmärksammade undersökning och citeringar via Google Scholar ger upphov till snedast fördelningar. Antalet publicerade arbeten ger upphov till de jämnaste. Likaså är det betydande omkastningar i vilka personer som hamnar i toppen mellan de olika måtten. Rangordningens resultat beror alltså i hög grad på vilket mått man använder.

Globaliseringen med tillhörande sammanlänkning av världens ekonomier har gjort det mycket enklare för forskarsamhället att föra ut resultat och knyta kontakter med forskare från andra länder än vad som tidigare varit fallet. Den ökade integrationen har gjort forskningsmarknaden större men också mer komplex; i fallet nationalekonomi finns i dag både en gemensam verktygslåda av generella nationalekonomiska begrepp och teorier och en rad nya världsomspännande men mer specialiserade underdiscipliner av ämnet. Av detta ökade samarbete och informationsflöde på den nationalekonomiska forskningsmarknaden följer ett kraftigt ökat behov av att utvärdera individer och institutioner, eftersom kostnaden av att ha personlig kännedom om deras kvalitet blir alltför hög. Givet detta blir det också viktigt att de mått vi använder verkligen mäter det vi vill att de ska mäta: forskningens kvalitet och vetenskapliga genomslag. Problemet är att dessa storheter inte kan mätas perfekt, och de mått som finns tillgängliga, såväl i nationalekonomi som i andra ämnen, är konstruerade av individer med egna syften och intressen, vilket riskerar att redan *a priori* gynna eller missgynna olika slags forskning. Inom andra vetenskaper, där erfarenheten av att rangordna forskningsprestationer har längre traditioner, pågår sedan länge en livlig debatt om styrkor och svagheter med olika mått (se t ex Hecht m fl 1998 och Cameron 2005), och vi menar att detta borde vara fallet också inom nationalekonomi.

Vad vill vi fånga i ett mått på nationalekonomisk forskningskvalitet? Citeringsbaserade mått är de i särklass vanligaste, men är god forskning nödvändigtvis den som citeras mest? Är tidskriftsartiklar det viktigaste sättet att föra ut forskningsresultat eller vill vi även ta med läroböcker, mono-

grafier och bokkapitel som mått på produktivitet och inflytande? Är det viktigaste i vilken tidskrift en artikel blivit publicerad eller hur många citeringar den får? Hur hanterar vi att vissa underdiscipliner publicerar många fler artiklar och därmed får fler citeringar? Hur värderar vi en forskare som publicerat en artikel med få citeringar i en högt rankad tidskrift jämfört med en forskare med flera välciterade artiklar i en lägre rankad specialtidskrift? Hur hanterar vi dynamiska problem, som att kvaliteten på olika tidskrifter och konkurrensen om att publicera förändras över tiden? Bryr vi oss om forskningsresultatens inflytande utanför den akademiska världen?[1]

De relativa vikterna av ovanstående frågor blir olika beroende på vad vi tänker använda resultatet av mätningen till, som t ex tjänstetillsättningar och anslagstilldelningar. Vår ambition är att undersöka om resultaten i hög grad påverkas av vilket mått vi väljer.

Det mått vi använder kommer också att ge signaler om vilken slags forskning som anses viktig och värdefull. Vi glömmer lätt invändningarna mot ett visst mått och kvar finns resultaten. Många har blivit medvetna om vad som mäts och det finns en stark tendens hos människor att ”göra det vi mäter” (Holmström och Milgrom 1991). Måttet kan också bli styrande för valet av forskningsproblem eller vart man väljer att skicka sin forskning för publicering.

En hel del ansträngningar har gjorts inom nationalekonomi för att mäta forskningsproduktion och rangordna forskare och forskningsinstitutioner. Först i USA (t ex Conroy och Dusansky 1995, Dusansky och Vernon 1998) och sedan i Europa (t ex Kalaitzidakis m fl 1999, 2003, Combes och Linnemer 2003 och Tombazos 2005). I Sverige har liknande mätningar varit mindre vanliga. Tobias Lindqvists undersökning i *Ekonomisk Debatt* från 2003 (Lindqvist 2003) var ett av de första mer systematiska försöken att mäta forskningsproduktion och rangordna forskare och institutioner/institut inom svensk nationalekonomi.[2]

Lindqvists undersökning har fått stor uppmärksamhet och även kritik (se t ex Sarafoglou 2003) trots att han ingående går igenom svagheterna med de mått han använder och själv betonar att resultaten därför bör tolkas med försiktighet.[3]

Vi vill med denna uppsats bredda diskussionen genom att studera vilka olika mått på forskningsproduktion som finns att tillgå, vilka svagheter och styrkor de olika måtten har, samt vilka resultat vi kan förvänta oss om vi använder oss av dem. Vi redovisar en egen empirisk undersökning av avidentifierad individuell forskningsproduktion inom svensk nationalekonomi där vi använder ett flertal olika mått på forskningsprestationer. Detta skiljer oss från Lindqvist som endast använder ett mått.

---

[1] Liknande frågor i en svensk nationalekonomisk forskningskontext diskuteras i Jonung (2005).

[2] Institutet för internationell ekonomi vid Stockholms universitet publicerade i slutet av 1990-talet listor över citeringsgrad bland svenska nationalekonomer.

[3] Även Kalaitzidakis m fl:s artiklar har väckt stor debatt, se t ex Tombazos (2005).

1. Vår undersökning

Vår population består av samtliga 88 professorer vid svenska universitet och högskolor födda 1940 eller senare.[4] Enbart de som inte uppnått folkpensionsålder har således tagits med. Alla redovisade resultat är avidentifierade eftersom syftet inte varit att göra ännu en rangordningslista av namngivna svenska forskare. Genomgående har verk med *n* författare som brukligt tilldelats 1/*n* i ”medförfattarvikt”. Perioden vi undersöker är 1986-juni 2005 och bestämdes av den kortaste tidsserien för de data vi ville undersöka. Informationen samlades in i juni-juli 2005.

Vår empiriska undersökning är en jämförelse av fem alternativa mått på forskningsprestation och i vilken rangordning vart och ett av dem resulterar när de tillämpas på svenska professorer i nationalekonomi. Måtten är följande:

a) KMS-vägda publikationer (Lindqvists mått)
b) *Impact Factor*-vägda publikationer
c) Antal citeringar av forskarens tre mest citerade verk i SSCI
d) Antal citeringar av forskarens tre mest citerade verk i Google Scholar
e) Total forskningsproduktion mätt som antal publicerade verk i *EconLit*

Vi har valt ut just dessa fem mått bland många andra tänkbara, eftersom de på olika sätt belyser de krav vi i inledningen menade kan ställas på mått som används för utvärdering av nationalekonomisk forskning.[5] Valet av vikter är mycket beroende av vilka och hur många tidskrifter som inkluderas och därmed ger publikationspoäng till forskaren. Olika mått kan ge upphov till stora skillnader mellan enskilda forskare, vilket avspeglar sig i stora variationer i de parvisa korrelationerna mellan olika mått. Coupé (2003) och Combes och Linnemer (2003) undersöker ett stort antal alternativa viktsystem utvecklade för t ex olika tidsperioder, populationer och där olika korrigeringar görs för bl a sidantal och teckentäthet. De senare finner att de parvisa korrelationerna varierar mellan 0,3 och 0,9. Gemensamt för alla dessa undersökta mått är att de avgränsar sig till citeringar mellan olika nationalekonomiska tidskrifter och oftast helt exkluderar citeringar från icke-ekonomisk litteratur. Kodrzycki och Yu (2005) visar att om icke-ekonomisk litteratur behandlades likvärdigt skulle betydande omkastningar ske i tidskriftsrangordningen.

Det mått som under senare år dominerat den europeiska nationalekonomiska diskussionen är det s k KMS-måttet (se nästa avsnitt). Detta mått ger en bra bild av nationalekonomisk forskning om man anser att publiceringar i ett fåtal men högt ansedda tidskrifter är det i särklass viktigaste. Dock kräver detta att man också accepterar KMS version av individuella tidskriftsvärderingar. Ett annat utvärderingsmått är *Impact Factor* eller IF-måttet, vilket sedan länge varit det dominerande utvärderingsmåttet inom naturvetenskaplig och medicinsk forskning. Detta mått gör inget försök att värdera

[4] Befordringsprofessorer är inkluderade. Individerna titulerades professor i juni 2005.
[5] Se t ex Coupé (2003) för en genomgång.

tidskrifters olika betydelse inom professionen utan viktar olika tidskrifter efter det antal citeringar de får. IF-måttet tar därmed inte hänsyn till att vissa underdiscipliner som t ex arbetsmarknads- och hälsoekonomi både citerar mer frikostigt och publicerar relativt fler artiklar. Ett annat problem är att IF-måttet fixeras för en tidsperiod på enbart två år (se vidare avsnittet om IF-måttet) och givet att det tar tid innan nya arbeten produceras och publiceras kommer måttet inte att fånga upp majoriteten av citeringarna till en artikel.

Både IF- och KMS-måtten antar att ett enskilt arbetes kvalitet är lika med kvaliteten hos en genomsnittlig artikel publicerad i den aktuella tidskriften genom att en artikels meritvärde är lika med tidskriftens historiska citeringsgrad. Detta visar sig ofta inte vara med verkligheten överensstämmande. Enbart hälften av de 18 artiklar som svenska professorer publicerat i *American Economic Review* (*AER*) hamnar bland respektive författares mest citerade artiklar. Hälften når inte heller ens upp till medelantalet citeringar för de artiklar vi undersöker, trots att *AER* är den i särklass mest citerade av alla tidskrifter.

Ett alternativt sätt att undersöka ett vetenskapligt inflytande som inte utgår från citeringsbaserade tidskriftsvikter (KMS och IF) är att räkna det faktiska antal citeringar som de enskilda publikationerna faktiskt får. En sådan analys är möjlig med hjälp av citeringsdatabaserna SSCI (*Social Sciences Citation Index*), som är en del av Thomson Scientifics välkända *Web of Science*, samt den internetbaserade sökmotorn Google Scholar (GS). Vi skapar ett mått för var och en av dessa databaser som innebär att vi summerar antalet citeringar (vägda för medförfattarskap) till en enskild forskares tre mest citerade verk enligt respektive databas. Även dessa mått har för- och nackdelar som vi diskuterar närmare nedan.

Slutligen har vi samlat in data på de enskilda forskarnas samtliga verk som finns upptagna i databasen *EconLit*, vilket indikerar att de publicerats i något internationellt sammanhang under den period vi undersöker. Anledningen till att vi inkluderar detta rena volymmått är att en stor del av nationalekonomisk forskning inte fångas upp av något av de fyra ovanstående måtten. De flesta tidskrifterna tilldelas inte poäng enligt KMS- och IF-systemen; många arbeten blir aldrig eller ytterst sällan citerade i SSCI eller Google Scholar. Detta gäller i synnerhet den vetenskapliga produktionen av böcker och bokkapitel.

KMS-vägda publikationer

Lindqvists artikel grundar sig på en tidskriftsrangordning utvecklad av Kalaitzidakis m fl (2003), hädanefter benämnd KMS.[6] Denna rangordning är en vidareutveckling av samma författares tidigare, mycket uppmärksammade försök att undersöka och värdera europeiska nationalekonomers publiceringar (Kalaitzidakis m fl 1999).

[6] Lindqvist använde *working paper*-versionen av KMS från 2002, där ingen hänsyn togs till antalet publicerade sidor per år vilket görs i KMS (2003), se tabell 1, kolumn 4 (Lindqvists serie) och 5 (sidantalsviktad).

Utgångspunkten för KMS är att kvalitetsrangordna tidskrifter, som ingår i SSCI, utifrån deras totala antal citeringar i andra tidskrifter, som också de ingår i SSCI, efter att vissa justeringar gjorts. Antalet citeringar publiceras för varje tidskrift och år i *Journal Citation Reports,* JCR (Thomson Scientific 2004). De justeringar KMS därefter gör är följande: först utesluts självciteringar (inom tidskriften) samt citeringar till ”äldre” artiklar (här: publicerade före 1994). Vidare tar KMS hänsyn till tidskriftens storlek genom att den viktas utifrån antalet artikelsidor i förhållande till genomsnittet för alla tidskrifter. Slutligen justerar KMS även för tidskrifternas ”allmänna inflytande på professionen” utifrån en särskild iterativ procedur som tar hänsyn till vilken status (dvs hur många citeringar) som de tidskrifter har från vilka citeringar av den undersökta tidskriften kommer. Med andra ord väger citeringar tyngre från de tidskrifter som själva är välciterade än från de som citeras sällan.[7] Slutresultatet är en rangordning av tidskrifter för perioden 1994–98 där var och en får en poäng som avspeglar dess inflytande på senare tid och där dubbelt så hög poäng motsvarar dubbelt så högt meritvärde.

Lindqvists undersökning baseras liksom vår på alla artiklar av svenska nationalekonomer som finns i artikeldatabasen *EconLit*.[8] *EconLit* innehåller ungefär 680 tidskrifter med pågående utgivning, samt ytterligare ca 175 vars utgivning numera upphört. Av dessa sammanlagt ca 855 tidskrifter ger enbart 141 poäng enligt KMS-måttet (eftersom det kräver att tidskriften även ingår i SSCI). Närmare 80 procent av de tidskrifter som finns med i *EconLit* ger inga poäng alls, hälften av det totala antalet poäng tillskrivs endast 10 tidskrifter och 2/3 av alla poäng som utdelas tillfaller de 20 högst poängsatta tidskrifterna. Översatt till poäng innebär KMS-måttet t ex att en artikel i *AER* är nästan lika mycket värd som femtio artiklar i *Journal of Health Economics*, medan samtliga 345 artiklar som någonsin publicerats i *Journal of Evolutionary Economics* enbart är värda drygt en halv artikel i *AER* eller en hel artikel i *Journal of Political Economy*.

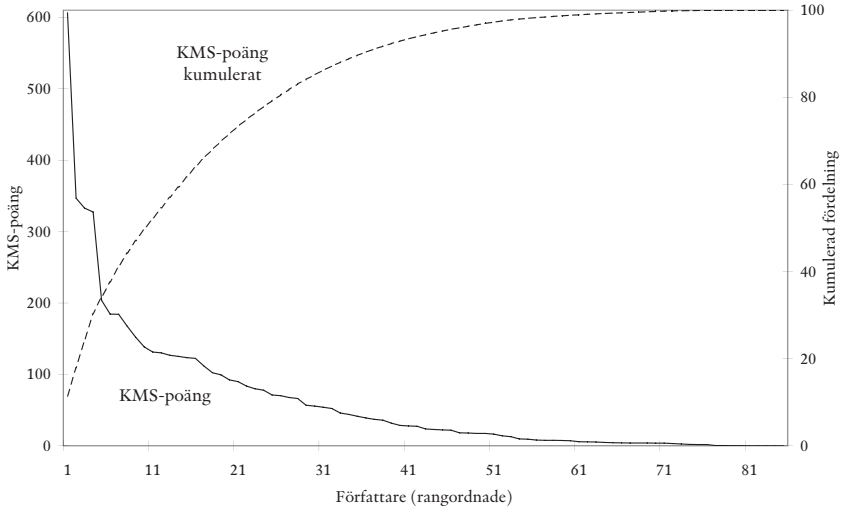
Den stora variationen i KMS-måttet mellan olika tidskrifter gav också stort genomslag på utfallet i Lindqvists rangordning. Tack vare en enda not i *AER* av en doktor från Södertörns Högskola erhåller Södertörns högskola nästan dubbelt så många poäng som det internationellt mycket framstående Beijerinstitutet och nästan 40 procent av ”värdet” av all den nationalekonomiska forskning som forskare verksamma vid SOFI vid Stockholms universitet publicerat sedan 1972 då institutet startade.

Nästa lika exceptionella effekter återfinns i rangordningen av enskilda forskare. Lindqvist visar att det fanns flera exempel på högproduktiva och välrenommerade forskare med en poängsumma som endast är i storleksordningen 10 procent av nivån för dem som ligger i toppskiktet.

---

[7] Palacios-Huerta och Volij (2004) anser även att citeringar från tidskrifter som i allmänhet citerar relativt lite bör ges större vikt än genomsnittet.

[8] *EconLit* innehåller ekonomiartiklar, brett definierat, som publicerats efter 1969. Databasen tillhandahålls av American Economic Association.



Figur 1
Svenska professorer i nationalekonomi rangordnade efter KMS-poäng och den kumulativa fördelningen av det totala antalet poäng i populationen

Anm: KMS-poängen (summan för forskarens artiklar efter att de multiplicerats med KMS-måttet för respektive tidskrift) avser publiceringar under perioden 1986-juni 2005. Samma poängskala tillämpas oavsett publiceringstidpunkt. Poängen har justerats för medförfattare. Populationen utgörs av professorer verksamma vid svenska universitet och högskolor vårterminen 2005 födda 1940 eller senare.

Trots att vår studie täcker en annan tidsperiod och population än Lindqvists, finner vi i princip samma fördelning av forskningsprestationerna när vi använder KMS-måttet.[9] På *x*-axeln i figur 1 finns de svenska nationalekonomer som ingår i studien, rangordnade från vänster till höger utifrån deras erhållna KMS-poäng. Diagrammet har två *y*-axlar: den vänstra mäter antal KMS-poäng för respektive person och den högra den kumulativa andelen av totalt erhållna KMS-poäng för hela populationen, vilken går mot 100 procent. Majoriteten av professorerna hamnar på förhållandevis försumbara poängnivåer; den undre halvan av populationen har tillsammans bara drygt fem procent av publiceringspoängen medan de tre högst rankade forskarna erhåller en fjärdedel av den totala publiceringspoängen. Med tanke på måttets konstruktion var resultatet väntat. Konkurrensen om att publicera sig i de tidskrifter som ger poäng i KMS-systemet är mycket hård och att ha upprepade publikationer i dem är endast ett fåtal förunnat.

Impact Factor-vägda publikationer

Medan KMS-måttet varit tongivande inom europeisk nationalekonomi är det mest kända viktsystemet inom andra vetenskapsområden, särskilt medicin och naturvetenskaper, det s k IF-måttet (*Impact Factor*). Detta är ett index som publiceras årligen i JCR, beräknat för varje tidskrift (i SSCI) och år genom att man dividerar alla årets citeringar av de två föregående årens artiklar med det totala antal artiklar som tidskriften publicerade

[9] Tidskriftsvikter baserade på KMS undersökning för perioden 1994–98.

under samma två år. Som ett exempel är 2005 års IF för tidskriften *AER* alla citeringar under 2005 av *AER*-artiklar publicerade 2003 och 2004 dividerat med samtliga artiklar som publicerades i *AER* under 2003 och 2004. Vi har använt IF-vikterna för år 2004 (Thomson Scientific 2004) och därmed antagit att dessa kan användas för att poängsätta tidskrifter under hela undersökningsperioden, dvs upp till 18 år tidigare.[10]

I jämförelse med KMS-måttet är IF-måttet liknande i så måtto att man kraftigt avgränsar vilka forskningsprestationer som räknas genom att enbart poängsätta artiklar i 160 tidskrifter. Däremot inkluderar IF-måttet självciteringar, vilket förstås är problematiskt, samt ger samma vikt till alla citeringar (medan KMS iterativa procedur viktar upp citeringar från relativt högt citerade tidskrifter).[11] Effekten av dessa båda skillnader mot KMS är förhållandevis stora omkastningar i hur tidskrifter värderas. Som tidigare nämnts krävs det enligt KMS-måttet nästan 50 artiklar i *Journal of Health Economics* för att uppnå motsvarande en *AER*-artikels poäng. Enligt IF-systemet är tvärtom en artikel i *Journal of Health Economics* mer värd än en artikel i *AER*, vilket förstås skulle kunna tyda på att vår föraning om att citeringsvanor inom olika underdiscipliner spelar roll för rankingen var riktig.[12] IF-måttet har sedan länge kritiserats för denna snedvridning, men även en lång rad andra problem som t ex det korta tvåårigа tidsfönstret för citerade artiklar.[13]

Som framgår av figur 2 ger också IF-metoden upphov till en statistiskt sett sned fördelning av poäng, om än i mindre utsträckning än KMS-metoden. Den undre halvan av populationen har ungefär 15 procent av totalpoängen. En knapp tredjedel av forskarna hamnar enligt IF-metoden på näst intill försumbara poängnivåer.

Antal citeringar av forskarens tre mest citerade verk i SSCI

Varken KMS eller IF-måtten säger någonting direkt om hur inflytelserik den individuella publiceringen är utan värderar den enbart utifrån hur inflytelserik den tidskrift är i vilken den publicerats. Om man i stället vill försöka värdera inflytandet hos ett enskilt arbete måste man undersöka i vilken omfattning det blivit citerat. Om enbart antalet citeringar räknas bygger detta alltså på antagandet att alla citeringar bör ges samma tyngd.

Vi har tagit fram ett par olika mått på citeringsgrad. Det första är sum-

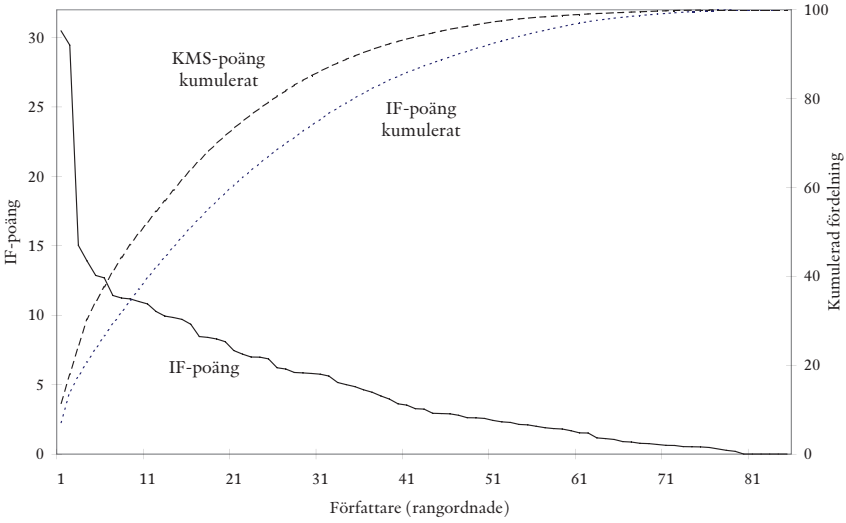
---

[10] Detta är givetvis en felkälla och egentligen skulle vi vilja använda olika IF-mått för olika publiceringsår. Dessa är dock inte, förvånande nog, enkelt tillgängliga bakåt i JCR. Denna felkälla spelar mindre roll just i vårt sammanhang där vi i först hand är ute efter att illustrera betydelsen av att använda olika mått och på vilket sätt valet av mått kan förändra värderingen av både själva forskningen och enskilda forskare.

[11] Här bör noteras att IF-måttet endast baseras på citeringar både till och från tidskrifter som ingår i Thomson Scientific.

[12] Enligt Thanasios Stengos (e-postkorrespondens), en av de forskare som utformat KMS-måttet, beror dock skillnaderna i viktning jämfört med IF framförallt på att rensning för självciteringar görs i det förra men inte det senare måttet.

[13] För mer om denna och annan kritik av IF-måttet, se t ex Hecht m fl (1998) och Cameron (2005).



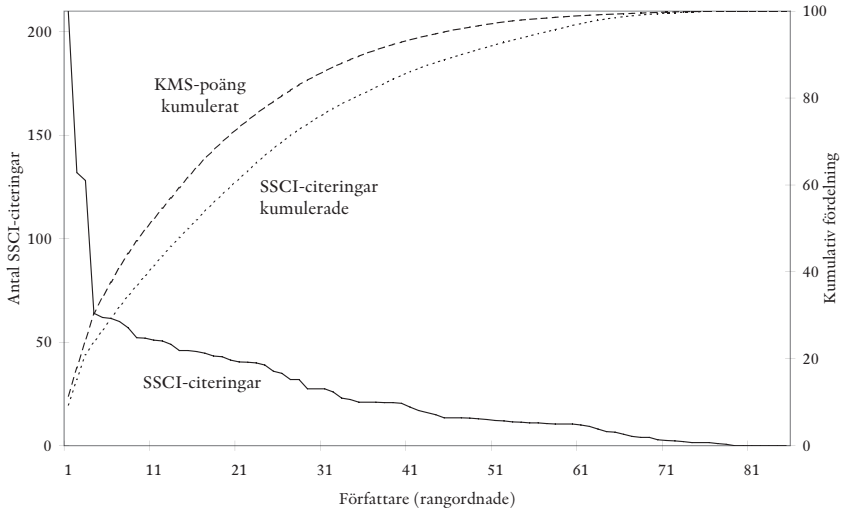
Figur 2
Svenska professorer i nationalekonomi rangordnade efter IF-poäng samt deras kumulativa fördelning över hela populationen

Anm: IF-poängen (summan för forskarens artiklar efter att de multiplicerats med IF-måttet för respektive tidskrift) avser publiceringar under perioden 1986-juni 2005 och samma poängskala tillämpas oavsett publiceringstidpunkt (IF-poäng för 2004). Poängen har justerats för medförfattare. Populationen utgörs av professorer verksamma vid svenska universitet och högskolor vårterminen 2005 födda 1940 eller senare.

man av citeringarna av varje professors tre mest citerade verk (topp 3-citeringarna) enligt databasen SSCI.[14] Vi har även här valt att justera för antalet medförfattare.[15] SSCI är den helt dominerande källan för citatredovisning, men systemet lider av ett antal brister av vilka några pekats ut av Holcombe (2004) och Klein (2005). Enbart citeringar i tidskrifter som ingår i Thomson Scientifics databaser registreras, och huruvida dessa är representativa för den totala citeringsmängden (dvs när även citeringar i icke-Thomson-tidskrifter, böcker mm räknas) är oklart.[16] Dessutom inkluderas självciteringar, vilket är problematiskt ur manipulationsperspektiv. Urvalsprocessen för vilka tidskrifter som ingår i SSCI har även beskrivits som oklar och bitvis inkonsekvent (Klein 2005). En annan nackdel med SSCIs citeringar

[14] Självfallet ligger det ett visst godtycke i att bara inkludera varje forskares tre mest citerade arbeten.
[15] Det är inte *a priori* givet att citeringar ska justeras för antal medförfattare. Forskning visar att meritvärdet för en samförfattad artikel är nästan lika högt som för en egenförfattad (Coupé 2003). Det är också tänkbart att forskare som rankas högt har lättare att attrahera framstående medförfattare och därmed har mer att vinna på om citeringar inte justeras för medförfattare. För att undersöka om så var fallet jämförde vi kvoterna mellan viktade och oviktade citeringar för tre olika urval av vår population: de tre respektive tio högst rankade forskarna och undre halvan av fördelningen. Om dessa kvoter var fallande skulle det antyda att framgångsrika forskare har lättare att få ökad hävstång genom att skriva med andra framgångsrika forskare. Resultaten är inte entydiga och vi väljer därför att följa praxis och justera antalet citeringar för delat författarskap.
[16] Det bör dock påpekas att citeringar av böcker eller tidskrifter utanför Thomsons databaser registreras.

Figur 3
Svenska professorer i nationalekonomi rangordnade efter summan av deras tre mest citerade arbeten enligt SSCI och den kumulativa fördelningen av det totala antalet poäng i populationen



Anm: Avser samtliga citeringar till arbeten publicerade under perioden 1986-juni 2005. Antalet citeringar har justerats för medförfattarskap. Populationen utgörs av professorer verksamma vid svenska universitet och högskolor vårterminen 2005 födda 1940 eller senare.

är att de länge enbart räknade det först angivna författarnamnet.[17] Detta har visserligen modifierats på senare tid, men det framgår inte alltid tydligt när justeringar skett (Lindqvist 2003). Särskilt i ljuset av att det finns tecken på att citatrangordningar fått större vikt på senare tid, t ex vid tjänstetillsättningar (Klein och Chiang 2004), är det viktigt att vara medveten om måttets brister.

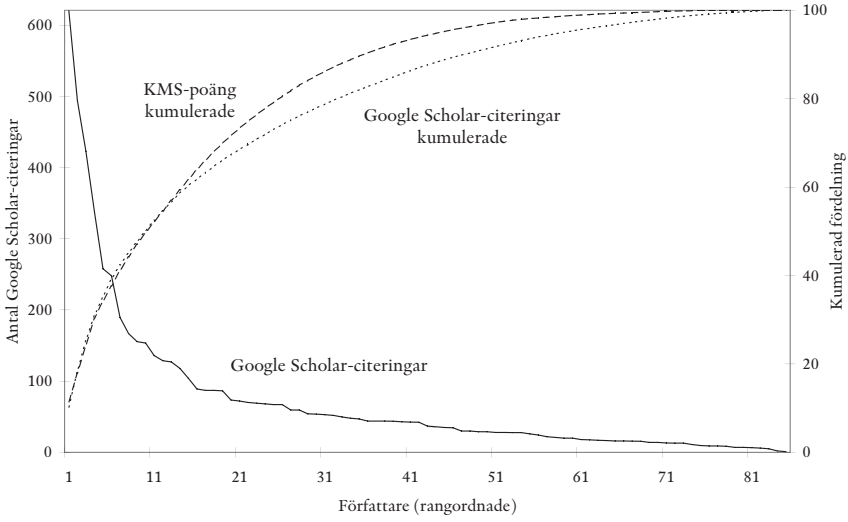
Som framgår av figur 3 är fördelningen av SSCI-citeringarna inte lika sned som KMS-poängen. Den lägst rankade tredjedelen får knappt fem procent av det totala antalet citeringar. Detta resultat överensstämmer väl med tidigare studier; enbart fem till tio procent av den forskning som produceras beräknas få ett stort inflytande på professionen, mätt som antal citeringar (Van Dalen och Klamer 2005). Medianantalet vägda citeringar för en svensk ekonomiprofessors mest citerade artikel är ungefär 14, vilket kan tyckas lågt. Det bör dock påpekas att 10 citeringar är att betrakta som mycket. Så många som 70 procent av alla publicerade artiklar har en eller ingen citering alls (Laband och Tollison 2003).

Antal citeringar av forskarens tre mest citerade verk i Google Scholar

Det andra måttet på individuella arbetens citeringar utgår från den relativt nyskapade sökmotorn Google Scholar (GS).[18] Till skillnad från SSCI

[17] Det finns också praktiska problem som engelsk stavning av svenska bokstäver, vilket vi tagit hänsyn till genom att söka på olika tänkbara stavningsvarianter.

[18] Google Scholar (http://scholar.google.com) är den mest välkända bland de nystartade internetbaserade sökmotorerna, även om Scopus (http://www.scopus.com) och Scirus (http://www.scirus.com) är andra nya alternativ.



Figur 4
Svenska professorer i nationalekonomi rangordnade efter summan av deras tre mest citerade arbeten enligt Google Scholar (GS) och den kumulativa fördelningen av det totala antalet citeringar i populationen

Anm: Urvalet utgörs av alla publikationer som kommer med i Google Scholar. Antalet citeringar har justerats för medförfattarskap. Populationen utgörs av professorer verksamma vid svenska universitet och högskolor vårterminen 2005 födda 1940 eller senare.

inkluderar GS citeringar till och från många olika typer av vetenskaplig produktion som publicerats på Internet, inklusive working papers, utredningar och t o m böcker, utöver tidskriftsartiklar. Om ett arbete både finns som working paper och publicerad artikel anger GS att dessa slås ihop så att dubbelräkning undviks (vi har dock observerat exempel på motsatsen). GS rensar emellertid inte för självciteringar, och har enligt vissa bedömare stora problem med täckningsgraden av tillgängliga publikationer (se t ex Jacsó 2005). Sökmotorns långsiktiga kvalitet och konsistens är ännu inte helt klarlagd varför våra resultat bör behandlas med viss försiktighet. Vi har valt att redovisa citeringsresultaten på samma sätt som för SSCI-analysen, dvs summan av topp 3-citeringarna.

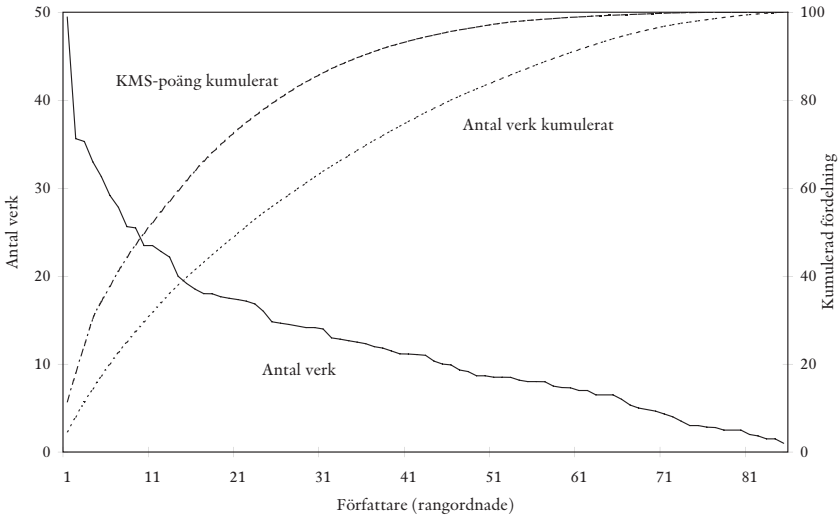
Som framgår av den vänstra *y*-axeln i figur 4 får de svenska ekonomiprofessorerna avsevärt fler citeringar när även andra verk än tidskriftsartiklar inkluderas. Fördelningen är något snedare än för SSCI-citeringarna, men mindre sned än fördelningen av KMS-poängen.

Total forskningsproduktion

Det sista måttet vi redovisar är antalet publicerade verk per person. Här avses enbart arbeten som finns upptagna i databasen *EconLit*. Vi har rensat bort samtliga working papers och återutgivna redan publicerade verk. Verk som förekommer flera gånger i databasen, exempelvis genom att de tryckts om i samlingsvolymer, räknas således bara en gång.

Enligt figur 5 varierar storleken på produktionen kraftigt i den svenska professorskåren. I genomsnitt har en svensk ekonomiprofessor producerat

Figur 5
Svenska professorer i nationalekonomi rangordnade efter antal verk i databasen *EconLit* och den kumulativa fördelningen av dessa publiceringar



Anm: Antalet verk har justerats för medförfattare. Populationen utgörs av professorer verksamma vid svenska universitet och högskolor vårterminen 2005 födda 1940 eller senare.

13 arbeten, men nästan en fjärdedel har under perioden publicerat färre än fem verk. Två tredjedelar av populationen står för 90 procent av produktionen.

2. Jämförelse av de olika måttens utfall

Låt oss nu sammanfatta på vilket sätt de redovisade metoderna ger olika bilder och i vilken utsträckning de kan sägas överlappa varandra. Vi väljer att både undersöka de forskare som hamnar i topp enligt respektive metod samt vilka bilder de olika måtten ger av produktionen för hela populationen av svenska nationalekonomiprofessorer. Detta redovisas i tabell 1.

Samtliga undersökta mått resulterar i sneda fördelningar; ett litet toppskikt står för huvuddelen av hela populationens poäng eller citeringar. Allra tydligast blir detta för KMS-måttet och de citeringar författarna får i GS. Enligt både KMS och GS bidrar de tre högst rankade forskarna med en fjärdedel av hela poängmassan, och när de tio högst rankade forskarna tagits med har hälften av poängen delats ut. Enligt KMS-metoden står den nedre halvan av populationen för enbart 6 procent av poängen. Om vi använder GS-citeringar som mått på forskningsproduktion hamnar två tredjedelar av populationen på försumbara poängnivåer. Detta bör kanske snarare ses som ett gott betyg till den mängd citeringar den högst rankade forskaren får än som ett bevis på att majoriteten av svenska nationalekonomiprofessorer har en helt försumbar forskningsproduktion. Om vi i stället väljer att studera antalet publicerade arbeten framstår produktiviteten i den svenska professorskåren i nationalekonomi som betydligt jämnare fördelad än vad

Tabell 1
Några jämförelser av utfallet för de fem använda måtten på forskningsproduktion, hela populationen

	Andel av totalsumman för resp mått i % för							
	Topp 3	Topp 10	Nedre halvan	Snitt topp 3 / median	Std.avv./ medel	Andel med försumbar nivå i %	Medel	Median
KMS (poäng)	24	50	6	18,0	1,6	30	62	27
IF (poäng)	17	37	13	7,7	1,1	11	5	3
SSCI (citeringar)	20	37	13	9,8	1,2	19	26	16
GS (citeringar)	25	50	13	13,6	1,5	67	70	37
Antal verk i EconLit	11	30	24	3,6	0,7	0	12	11

Anm: Med ”försumbar nivå” avses här mindre än 1 procent av den högsta poängen enligt respektive mått.

KMS-metoden och GS-citeringarna indikerar; en fjärdedel av antalet verk har publicerats av den nedre halvan i fördelningen. IF-metoden och antalet SSCI-citeringar ger bilder som ligger mellan dessa två ytterligheter.

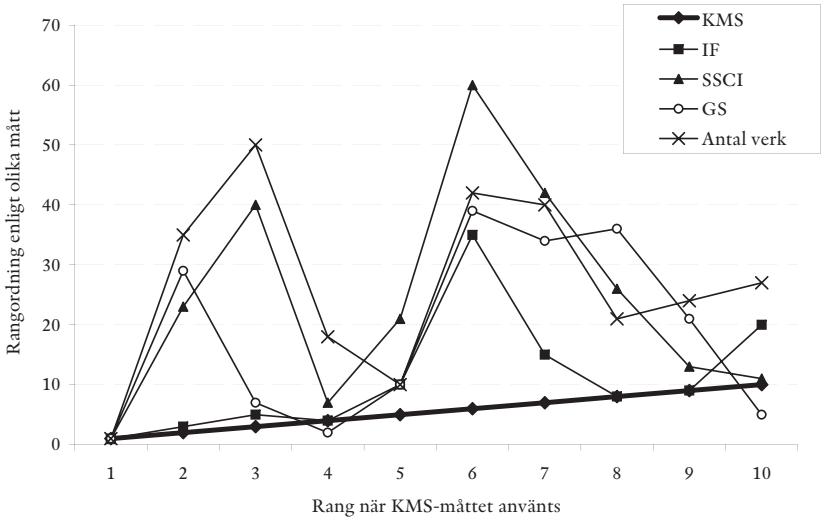
Vilka forskare når toppen?

Vilka individer som återfinns i det översta skiktet växlar mellan metoderna, och vilket mått vi använder för att rangordna ekonomer kommer alltså att direkt påverka vilka individer vi återfinner såväl på toppositioner som längre ner i rangordningen.

För KMS- och IF-måtten finns dock betydande överlappningar; sex av de tio främsta enligt KMS återfinns bland topp 10 även enligt IF. De professorer som återfinns i båda måttens topp 10-grupp verkar inom skilda fält, men spridningen i detta hänseende får anses vara klart större för IF-måttet. Bland dess tio främsta återfinns både renodlade teoretiker, renodlade empiriker och forskare som är utpräglat förknippade med ett visst specialområde. Fyra av de tio främsta enligt IF-metoden hör dessutom till de mest produktiva mätt i antal publicerade verk. Kopplingen mellan storlek på produktionen mätt som antalet publicerade verk och antalet KMS-poäng är mindre tydlig; av de tio högst rankade forskarna enligt KMS-måttet hör endast två till de tio mest produktiva mätt i antal publicerade arbeten.

Vad gäller SSCI-citeringar står de tre mest citerade professorerna för en femtedel av det totala antalet citeringar av varje professors tre mest citerade arbeten. Det är inte de forskare som rankas högst enligt KMS- eller IF-måtten som är de mest citerade; enbart två av de mest citerade hör till toppskiktet enligt KMS respektive tre för IF. Det är inte heller de som publicerar mest som citeras mest; enbart två av de tio mest citerade hör till de tio mest produktiva mätt i antal publikationer. Här bör noteras att om vi i stället mätt antalet citeringar till samtliga arbeten så kunde bilden varit

Figur 6
Rangen enligt de övriga fyra måtten för de tio högst rankade forskarna enligt KMS-metoden



annorlunda. Även GS-citeringarna ger en i hög grad annorlunda bild än övriga redovisade metoder. Endast två à tre av de tio mest citerade enligt GS återfinns bland de tio främsta enligt de andra metoderna.

Att överlappningen mellan de tio forskare som rankas högst enligt KMS är långt ifrån fullständig framgår också tydligt av figur 6.[19] Flera av forskarna från topp 10-listan rankas betydligt sämre enligt de andra måtten. Exempelvis hör den som rankas som nummer 6 enligt KMS inte ens till de 30 högst rankade forskarna enligt någon av de andra metoderna.

Grad av överlappning i hela populationen

För att undersöka i vilken utsträckning de olika måtten kan sägas ge samma bild i hela populationen har vi beräknat parvisa Pearsonkorrelationer för de fem olika måtten på forskningsproduktion baserat på forskarnas respektive rang. Dessa presenteras i tabell 2.

Graden av överensstämmelse mellan måtten bekräftar det vi kom fram till då vi jämförde utfallet för toppskiktet enligt respektive metod. KMS och det antal arbeten en forskare producerat visar minst grad av överensstämmelse. KMS och IF är de mått som ger mest likartade utfall. De olika citeringsmåtten är ungefär lika starkt korrelerade med IF-måttet, medan KMS-måttet däremot i betydligt högre grad är korrelerat med GS-citeringar än med SSCI-citeringar. Detta var oväntat för oss med tanke på att en författare erhåller KMS-poäng från publikationer inom ett förhållandevis snävt urval av tidskrifter, medan GS mäter citeringar inom en bred definition av publikationer och även inflytande utanför den akademiska världen. Bland de allra mest citerade verken på GS återfinns t ex två läroböcker och ett par working papers. Endast tre av de tio mest citerade verken har publicerats i topprankade tidskrifter.

[19] Notera att detta inte är samma forskare som i Lindqvist (2003) pga att populationen inte är densamma.

Tabell 2
Korrelationsmatris för de olika måtten på forskningsproduktivitet, hela populationen

	KMS	IF	SSCI	GS	Antal verk
KMS	1				
IF	0,78	1			
SSCI	0,56	0,63	1		
GS	0,70	0,66	0,62	1	
Antal verk	0,50	0,72	0,52	0,54	1

Anm: Samtliga koefficienter är signifikant skilda från noll på 5-procentsnivån.

3. Resultatens känslighet

Vi har visat att oavsett vilket mått vi använder får vi en bild av svensk nationalekonomisk forskning som indikerar att produktiviteten och prestationerna är skevt fördelade. Givet detta är det intressant att undersöka hur känsliga resultaten är för extrema observationer, dvs vad som händer om vi exkluderar t ex de tre högst rankade forskarna enligt respektive mått från populationen.

När de tre högst rankade forskarna uteslutits kvarstår många av slutsatserna från analysen av hela populationen. De olika måtten ger fortfarande upphov till sneda fördelningar, men spridningen i fördelningen har minskat. De tre främstas andel av den totala poängen har minskat med ungefär en tredjedel. Det är också intressant att notera att vi får en relativt sett större förändring av de mått som mäter citeringsgrad än av KMS- och IF-metoderna, vilket antyder att de förstnämnda metoderna verkar vara mer känsliga för extremvärden pga större spridning i data. Sammanfattningsvis kan denna känslighetsanalys och studiet av hela populationen sägas ge vid handen att toppskiktet i svensk nationalekonomi är större än tre personer: knappt tio procent utgör en tätgrupp och deras prestationer bedöms stå i en viss särklass oavsett vilket mått som används. Detta resultat ligger i linje med vad andra studier funnit för internationell nationalekonomisk forskning (se t ex Coupé 2003).

När vi åter beräknar parvisa korrelationer finner vi att korrelationerna mellan de olika måtten i de flesta fall sjunker, vilket tyder på att måtten framför allt stämmer överens när det gäller de forskare som rankas högst. Korrelationen mellan KMS- och IF-måtten är identisk när vi jämför hela populationen med det stickprov där de tre mest produktiva forskarna exkluderats. Detta indikerar att överensstämmelsen mellan dessa metoder inte enbart drivs av att de ger samma utfall för de högst rankade forskarna.

4. Slutsatser

Med ökad specialisering och internationalisering av forskningen har behovet av jämförelser mellan forskare och forskningsinstitutioner ökat markant. Det finns trots detta ännu ingen konsensus om hur sådana rangordningar

ska genomföras och vilka mått på forskningskvalitet som ska användas. Vi har i denna artikel presenterat rangordningar av svenska nationalekonomiprofessorers forskningsprestationer baserat på fem väletablerade mått på forskningsprestationer. Samtliga mått har sina för- och nackdelar och det är i de flesta sammanhang inte givet vilket av dem som är det mest naturliga att använda. Vår undersökning visar att rangordningsutfallen varierar kraftigt mellan måtten, men att detta är delvis förutsägbart utifrån hur respektive mått är konstruerat och dess betoning på värdet av olika typer av publikationer. Detta antyder att det är mycket viktigt att på förhand diskutera vilka mått som kan och bör användas i en rangordning eftersom utfallet är i så hög grad beroende av vilket mått som används.

Mer specifikt visar vår jämförelse att Google Scholar och KMS är de mått som ger högst koncentration i toppen av fördelningarna. Den nedre halvan av de svenska professorerna har endast hälften så stor andel av det totala antalet poäng enligt KMS-måttet jämfört med de andra citeringsmåtten (IF, SSCI och Google Scholar) och endast en fjärdedel så stor som andelen av det totala antalet publicerade verk (sex respektive 24 procent). Dessutom får toppen av fördelningen (antingen topp 3 eller topp 10) konsekvent störst andel av totalproduktionen när KMS- eller Google Scholar-måttet används, även om skillnaderna här är något mindre än vid en jämförelse av den nedre halvan av fördelningen. Det bör dock understrykas att även om fördelningarna liknar varandra är det till en del fråga om olika individer som positioneras i toppen respektive botten med de båda måtten.

Hur en forskare rangordnas i förhållande till sina kolleger spelar en stor roll för möjligheterna att få anslag och för avancemang i organisationer. Det är därför inte bara viktigt att vara medveten om att ett enskilt mått kan ge en vinklad bild när två enskilda forskare jämförs, utan att oavsett vilket eller vilka mått som än används så kommer det att ha stor inverkan på hur forskare arbetar och vilka frågor de väljer att forska om. Användningen av ett visst jämförelsemått får en avgörande betydelse för hela incitamentsstrukturen och därmed också för hela forskningsmiljön (Holcombe 2004). Denna effekt kan också förväntas växa över tiden när makten över att definiera vad som är forskning av hög kvalitet successivt flyttas över till forskare som redan från början av den egna karriären skolats in i att framgång mäts i termer av vissa kriterier.

Cameron, B D (2005), ”Trends in the Usage of ISI Bibliometric Data: Uses, Abuses, and Implications”, *Libraries and the Academy*, vol 5, s 105-125.

Combes, P-P och L Linnemer (2003), ”Where Are the Economists Who Publish?”, *Journal of the European Economic Association*, vol 1, s 1250-1308.

Conroy, M och R Dusansky (1995), ”The Productivity of Economics Departments in the US: Publications in the Core Journal”, *Journal of Economic Literature*, vol 33, s 1966-1971.

Coupé, T (2003), ”Revealed Performances: Worldwide Rankings of Economists and Economics Departments, 1990-2000”, *Journal of the European Economic Association*, vol 1, s 1309-1345.

Dusansky, R och C J Vernon (1998), ”Rankings of U.S. Economics Departments”, *Journal of Economic Perspectives*, vol 12, s 157-170.

Hecht F, B K Hecht och A A Sandberg (1998), ”The Journal ’Impact Factor’: A Misnamed, Misleading, Misused Measure”, *Cancer Genet Cytogenet*, vol 104, s 77-81.

Holcombe, R G (2004), ”The National Research Council Ranking of Research Universities: Its Impact on Research in Economics”, *Econ Journal Watch*, vol 1, s 498-514.

Holmström, B och P M Milgrom (1991), ”Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design”, *Journal of Law, Economics and Organization*, vol 7, s 24-52.

Jacsó, P (2005), ”Google Scholar: the Pros and Cons”, *Online Information Review*, vol 29, s 208-214.

Jonung, L (2005), ”Vad bör en nationalekonom göra?”, stencil, Europakommissionen, Bryssel.

Kalaitzidakis, P, T Mamuneas och T Stengos (1999), ”European Economics: An Analysis Based on Publications in the Core Journals”, *European Economic Review*, vol 43, s 1150-1168.

Kalaitzidakis, P, T Mamuneas och T Stengos (2003), ”Rankings of Academic Journals and Institutions in Economics”, *Journal of the European Economic Association*, vol 1, s 1346-1366.

Klein D B (2005), ”The Social Sciences Citation Index: A Black Box – with an Ideological Bias?”, *Econ Journal Watch*, vol 1, s 134-165.

Klein D B och E Chiang (2004), ”Citation Counts and SSCI in Personnel Decisions: A Survey of Economics Departments”, *Econ Journal Watch*, vol 1, s 166-174.

Kodrzycki, Y K och P D Yu (2005), ”New Approaches to Ranking Economics Journals”, Working Paper 05-12, Federal Reserve Bank of Boston, Boston.

Laband, D och R Tollison (2003), ”Dry Holes in Economic Research”, *Kyklos*, vol 56, s 161-174.

Lindqvist, T (2003), ”Nationalekonomisk forskning i Sverige – publiceringar och rankning av forskare och institutioner”, *Ekonomisk Debatt*, årg, 31, nr, 3, s 21-30.

Palacios-Huerta, I och O Volij (2004), ”The Measurement of Intellectual Influence”, *Econometrica*, vol 72, s 963-977.

Sarafoglou, N (2003), ”Diskutabel ranking – kommentar till Lindqvist”, *Ekonomisk Debatt*, årg 31, nr 3, s 39-42.

Thomson Scientific (2004), *Journal Citation Reports 2004. Social Sciences Edition*, The Institute for Scientific Information, Philadelphia.

Tombazos, C (2005), ”A Revisionist Perspective of European Research in Economics”, *European Economic Review*, vol 49, s 251-277.

Van Dalen, H och A Klamer (2005), ”Is Science A Case of Wasteful Competition?”, *Kyklos*, vol 58, s 395-414.