



IFAU

Institutet för arbetsmarknads- och
utbildningspolitisk utvärdering

Att skatta intergenerationella inkomstsamband: en jämförelse av de vanligaste måtten

**Martin Nybom
Jan Stuhler**

RAPPORT 2015:9

Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU) är ett forskningsinstitut under Arbetsmarknadsdepartementet med säte i Uppsala. IFAU ska främja, stödja och genomföra vetenskapliga utvärderingar. Uppdraget omfattar: effekter av arbetsmarknads- och utbildningspolitik, arbetsmarknadens funktionssätt och arbetsmarknadseffekter av socialförsäkringen. IFAU ska även sprida sina resultat så att de blir tillgängliga för olika intressenter i Sverige och utomlands.

IFAU delar även ut forskningsbidrag till projekt som rör forskning inom dess verksamhetsområden. Forskningsbidragen delas ut en gång per år och sista dag för ansökan är den 1 oktober. Eftersom forskarna vid IFAU till övervägande del är nationalekonomer, ser vi gärna att forskare från andra discipliner ansöker om forskningsbidrag.

IFAU leds av en generaldirektör. Vid institutet finns ett vetenskapligt råd bestående av en ordförande, institutets chef och fem andra ledamöter. Det vetenskapliga rådet har bl.a. som uppgift att lämna förslag till beslut vid beviljandet av forskningsbidrag. Till institutet är även en referensgrupp knuten där arbetsgivar- och arbetstagersidan samt berörda departement och myndigheter finns representerade.

Rapporterna finns även i tryckt format. Du kan beställa de tryckta rapporterna via telefon eller mejl. Se nedanstående kontaktinformation.

Postadress: Box 513, 751 20 Uppsala
Besöksadress: Kyrkogårdsgatan 6, Uppsala
Telefon: 018-471 70 70
Fax: 018-471 70 71
ifau@ifau.uu.se
www.ifau.se

IFAU har som policy att en uppsats, innan den publiceras i rapportserien, ska seminariebehandlas vid IFAU och minst ett annat akademiskt forum samt granskas av en extern och en intern disputerad forskare. Uppsatsen behöver dock inte ha genomgått sedvanlig granskning inför publicering i vetenskaplig tidskrift. Syftet med rapportserien är att ge den ekonomiska politiken och den ekonomisk-politiska diskussionen ett kunskapsunderlag.

Att skatta intergenerationella inkomstsamband: en jämförelse av de vanligaste måtten^a

av

Martin Nybom^b och Jan Stuhler^c

2015-06-16

Sammanfattning

Ett växande forskningsfält studerar i vilken utsträckning skillnader i ekonomisk status fortlever mellan generationer. Ofta används begreppet *intergenerationell rörlighet*, vilket knyter an till social rörlighet och utjämning av livschanser. Syftet är ofta att beskriva olika statistiska samband mellan barns och föräldrars livsinkomster, där starka samband innebär låg rörlighet och att barn i hög utsträckning ”ärver” sina föräldrars inkomster (och vice versa). Att undersöka livsinkomster ställer dock stora krav på data, och i stället används ofta mer kortsiktiga mått, såsom årsinkomster. Det vanligaste rörlighetsmålet, den intergenerationella elasticiteten, kan dock bli gravt felskattat när man låter årsinkomster approximera livsinkomst. Dock vet vi lite om graden av felskattning för andra rörlighetsmått. Vi använder inkomstdata för åren 1960–2007, kopplade till fäder och deras söner, för att undersöka detta för de fyra vanligaste måtten. Mått baserade på rangordningar i inkomstfördelningen tycks vara betydligt mer stabila och pålitliga än de traditionella måtten. I vår analys finner vi att rörligheten i Sverige *i genomsnitt* är relativt hög, men mycket lägre bland de allra fattigaste och rikaste i befolkningen. Den intergenerationella rörligheten är som mest felskattad för just dessa grupper, vilket tenderar att dölja att det både för de allra fattigaste och allra rikaste finns ett starkt samband mellan fäders och söners ekonomiska situation.

^a Denna rapport är en sammanfattning av ”Biases in standard measures of intergenerational income dependence”, IFAU Working paper 2015:13.

^b SOFI, Stockholms universitet, och IFAU. martin.nybom@sofi.su.se

^c Universidad Carlos III Madrid, och SOFI, Stockholm universitet. jstuhler@eco.uc3m.es

Innehållsförteckning

1	Inledning	3
2	Data.....	5
3	Metod och grundläggande resultat.....	5
4	Fyra vanliga mått på intergenerationella samband i inkomster	10
4.1	Elasticiteten	10
4.2	Den linjära korrelationen	10
4.3	Rangkorrelationen	14
4.4	Transitionsmatriser och kopulan av barns och föräldrars inkomster	14
5	Slutsatser.....	17
	Referenser	19

1 Inledning

Ett växande forskningsfält studerar till vilken grad skillnader i socioekonomisk status överförs mellan generationer. Inom nationalekonomin används ofta begreppet *intergenerationell rörlighet*, och fokus ligger vanligtvis på i vilken utsträckning barn ”ärver” sina föräldrars inkomster. Sociologer talar i stället ofta om *social rörlighet*, med huvudsaklig inriktning på individers klasstillhörighet eller yrkesstatus. Båda dessa studieområden relaterar till den mer allmänna diskussionen om ojämlikhet i livschanser och det brett omhuldade målet att ge individer ”lika förutsättningar”.

Forskare med fokus på inkomstdimensionen söker vanligtvis att empiriskt skatta statistiska samband mellan föräldrars och barns livsinkomster. Ett starkt samband innebär att föräldrars ekonomiska situation i hög grad överförs till deras barn, och att samhället uppvisar en låg grad av intergenerationell rörlighet (och vice versa). Begreppet livsinkomst är här ytterst centralt, och kan ses som antingen summan eller genomsnittet av alla årsinkomster under arbetslivet. Dock ställer det mycket stora krav på data att ha sådana uppgifter för två generationer, och i stället används i praktiken ofta betydligt mer kortsiktiga eller temporära mått, såsom årsinkomster.

Det vanligaste måttet för att skatta intergenerationell rörlighet, den intergenerationella *elasticiteten*, anger med hur många procent ett barns inkomst förväntas avvika från genomsnittet i sin generations inkomstfördelning baserat på sin förälders avvikelse från genomsnittet i föräldrarnas inkomstfördelning. Detta mått kan dock bli rejält felskattat om temporära inkomstobservationer, såsom årsinkomster, används för att approximera livsinkomster. En ökad förståelse för de statistiska orsakerna till dessa felskattningar har lett fram till slutsatsen att den intergenerationella rörligheten i många fall har varit överskattad, och bidragit till en kraftigt reviderad bild av hur socialt och ekonomiskt rörliga många länder i den industrialiserade världen är.¹

Delvis som en motreaktion så har forskningen i ökad utsträckning börjat använda sig av andra rörlighetsmått: *linjära korrelationer*, vilka till skillnad från elasticiteten tar hänsyn till att inkomstspridningen kan förändras mellan generationer; *rangkorrelationer*, vilka fångar sambandet mellan föräldrars och barns relativa positioner (dvs. rangordningar) i respektive inkomstfördelning; och *rangbaserade transitionsmatriser*, ofta med fokus på de allra fattigaste eller rikaste i befolkningen. Rangbaserade mått fokuserar på individers position i inkomstfördelningen, ofta uttryckt som inkomstpercentiler. Enkelt uttryckt har en

¹ Det finns ett flertal läsvärda litteraturöversikter på detta område. Se till exempel Solon (1999), Björklund & Jäntti (2009), Black & Devereux (2011), och Jäntti & Jenkins (2014).

person på 10:e inkomstpercentilen högre inkomst än 10 procent av befolkningen (och lägre inkomst än 90 procent), medan en person på 50:e percentilen befinner sig mitt i inkomstfördelningen. Att konstruera inkomstfördelningen på detta sätt innebär att man använder sig av en identisk fördelning för både olika länder och för olika tidsperioder och generationer inom länder. Dessa typer av mått utgör också grunden till många av den senaste tidens studier av skillnader i rörlighet mellan länder (Corak m.fl., 2014; Bratberg m.fl., 2015), över tid (Chetty m.fl., 2014b; Pekkarinen m.fl., 2015) och mellan regioner inom länder (Chetty m.fl., 2014a).

Tyvärr vet vi dock relativt lite om hur robusta dessa alternativa rörlighetsmått är. Måtten har ännu inte blivit systematiskt analyserade med hjälp av ”riktig” data på livsinkomster.

I den här rapporten visar vi hur graden av felskattning skiljer sig åt mellan de fyra måtten på intergenerationell inkomströrlighet som nämnts ovan. Vår empiriska analys drar nytta av de ovanligt långa registerserierna av inkomstdata som finns tillgängliga i Sverige. Att vi observerar nästan hela yrkeskarriären för såväl föräldrar som deras barn gör att vi kan ta fram benchmark-skattningar (dvs. i princip ”sanna” samband) och jämföra dessa med skattningar baserade på årsinkomster vid olika åldrar. Huvudfokus i denna rapport ligger på de direkta empiriska resultaten. För en mer detaljerad diskussion av effekterna av mätfel på de olika rörlighetsmåtten hänvisar vi till vårt IFAU working paper (Nybom & Stuhler, 2015, på engelska).

Vår analys visar att effekterna av dessa typer av mätfel i livsinkomster varierar stort mellan olika rörlighetsmått. De observerade mönstren är dock relativt systematiska och generaliserbara. Ett viktigt resultat är att rangbaserade mått i lägre grad än traditionella mått överskattar den intergenerationella rörligheten och är mer stabila över ålder. Användandet av sådana mått har därför möjligheten att motverka mycket av den mätproblematik som för tillfället hämmar detta forskningsfält. Vår analys av olika delar av inkomstfördelningen avslöjar också ett par intressanta mönster av relevans för den svenska diskussionen. Medan rörligheten i Sverige *i genomsnitt* är relativt hög, och i linje med tidigare svenska studier, så finner vi att den är diametralt annorlunda och mycket lägre bland de allra fattigaste och rikaste i befolkningen. Det är därför noterbart att skattningar baserade på temporära inkomster för just dessa grupper också tenderar att vara de mest felskattade, vilket i sin tur tenderar att dölja dessa skillnader.

I nästa avsnitt beskriver vi de data vi använder och i avsnitt 3 sammanfattar vi våra grundläggande empiriska resultat. Avsnitt 4 innehåller en mer detaljerad diskussion av de respektive måtten och ytterligare empiriska resultat. Vi

studerar även hela fördelningen av föräldrars och deras barns inkomster ("kopulan") för att undersöka om resultaten i genomsnitt skiljer sig åt för olika inkomstgrupper. Vi presenterar våra slutsatser i avsnitt 5.

2 Data

Vi använder data från två dataregister, sammanlänkade av Statistiska centralbyrån: flergenerationsregistret, som ger oss ett representativt urval av den svenska befolkningen, inklusive information om respektive individs biologiska föräldrar; och inkomstregistret, från vilket vi använder total förvärvsinkomst baserat på inkomstskattedeclarationer. Vi inkluderar söner födda 1953–1957 med fäder födda 1927–1941. Inkomstdata finns tillgängliga för åren 1960–2007 och vi begränsar vårt urval till fäder och söner med registrerad inkomst under åtminstone tio olika år.

Vårt huvudurval består av 6 525 par av fäder och söner, med söners inkomster observerade mellan 22 och 50 års ålder, och fäders inkomster observerade mellan 33 och 60 års ålder, i båda fallen oberoende av sonens och faderns respektive födelseår. Vi dividerar summan av reala årsinkomster med antalet årliga inkomstobservationer per individ. För våra analyser av linjära korrelationer standardiserar vi års- respektive livsinkomster baserat på födelseår. För våra rangbaserade analyser så rangordnar vi individer baserat på (absoluta) inkomstnivåer.

För att analysera hela fördelningen av föräldrars och deras barns inkomster använder vi ett utökad urval för vilket vi begränsar de observerade inkomståren till 1968–2007. Den huvudsakliga skillnaden med detta urval är att fäders inkomster nu observeras från 41 till 60 års ålder. Alla övriga urvalskriterier och variabeldefinitioner är oförändrade. Detta utökade urval består av 63 441 par av fäder och söner. På det stora hela erbjuder dessa data en unik möjlighet att studera olika mått på intergenerationell rörlighet baserat på närapå livslånga inkomstdataserier för två generationer.

3 Metod och grundläggande resultat

Vi kan benämna föräldrars och barns livsinkomster² x^* och y^* . Sambandet mellan y^* och x^* kan sammanfattas med hjälp av olika mått, såsom den intergenerationella elasticiteten (lutningskoefficienten i en regression av y^* på x^*),

² I praktiken används logaritmerade livsinkomster (eller årsinkomster) när analysen görs. För enkelhetens skull skrivs detta dock inte ut i texten.

linjära- eller rangkorrelationer, eller rangbaserade transitionssannolikheter (förklaras mer i detalj nedan). I empiriska studier så observeras dock vanligtvis endast inkomster för en kortare period (t.ex. årsinkomster). Dessa kan skrivas som $x = x^* + u$ och $y = y^* + v$, där u och v kan betraktas som approximerings- eller *mätfel*. Exempel på varför mätfel uppstår är att t.ex. inkomstobservationer under studietiden eller inkomsten under en period av arbetslöshet tenderar att underskatta livsinkomsten.

När man använder sig av kortsiktiga inkomstmått i stället för sanna livsinkomster kan skattningen av det intergenerationella sambandet därför uppvisa ett så kallat "bias". Ett "bias" är en form av felskattning, vilken innebär att den statistiska analysen ger resultat som systematiskt skiljer sig från det sanna samband som man vill fånga. När det gäller intergenerationella samband kan skattningarna skilja sig från det sanna sambandet av två anledningar; det finns alltså två huvudsakliga orsaker till att den intergenerationella rörligheten felskattas. För det första, slumpmässiga (eller "klassiska") mätfel kan ge ett "*bias mot noll*", vilket innebär att det statistiska sambandet underskattas. Utöver detta så kan heterogenitet i inkomstprofiler (över ålder) göra så att sambanden mellan mätfelen och de sanna värdena varierar systematiskt över ålder, vilket skapar ett s.k. *livscykelbias* (Jenkins, 1987). Ett exempel på detta är att högutbildade, vilka tenderar att ha höga livsinkomster, vanligtvis har relativt låga inkomster vid unga åldrar men sedan betydligt högre inkomstökningar under arbetslivet. Detta innebär att inkomstskillnader i termer av livsinkomster tenderar att underskattas om man tittar på yngre personer och överskattas om man tittar på äldre personer.

Vår metod är enkel. Vi jämför benchmark-skattningar som baseras på observerade långsiktiga inkomster (dvs. y^* och x^*) med ålderbaserade skattningar baserade på årsinkomster för en eller båda generationerna. Vi fokuserar först på mätfel i sönernas livsinkomst y^* , genom att använda årsinkomster för söner men livsinkomster för fäder. Figur 1 visar våra grundläggande empiriska resultat. För vart och ett av de fyra olika rörlighetsmåten visas benchmark-skattningar (heldragen linje) och åldersspecifika skattningar baserade på årsinkomster för barnen (streckad linje). Benchmark-skattningarna visar alltså de "sanna" intergenerationella sambanden, och avvikelsen mellan dessa och de åldersspecifika skattningarna visar i vilken grad analyser baserade på årsinkomster leder till felskattningar.

Genomgående för samtliga rörlighetsmått är att de underskattar de intergenerationella sambanden stort vid unga åldrar. Då individer med höga livsinkomster tenderar ha låga inkomster vid unga åldrar, till exempel på grund av längre skolgång, är varken linjära- eller rangbaserade mått särskilt nära sina

respektive benchmarks innan 30-årsåldern. Vid senare åldrar varierar dock både graden av och riktningen på felskattningarna stort.

Elasticitetsskattningar är tydligast påverkade av vad som brukar kallas livscykeffekter: de ökar närmast monotont med ålder och överskattar det intergenerationella sambandet vid högre åldrar. I linje med tidigare studier så är dessa livscykeffekter minst när inkomster observeras mitt i karriären (se Grawe, 2006; Haider and Solon, 2006). Variationen i årsinkomstbaserade skattningar är dock betydande; mellan 30 och 50 års ålder dubblas estimaten från cirka 0,2 till 0,4. Både linjära- och rangkorrelationen är betydligt mindre känsliga för sådana livscykeffekter. Den linjära korrelationen ökar endast marginellt mellan 30 och 50 års ålder, medan rangkorrelationen är förvånansvärt stabil över samma intervall. Att den senare faller marginellt i sena 40-årsåldern tyder dock på att även rangkorrelationen kan bli mindre pålitlig då inkomster observeras i slutet av karriären.

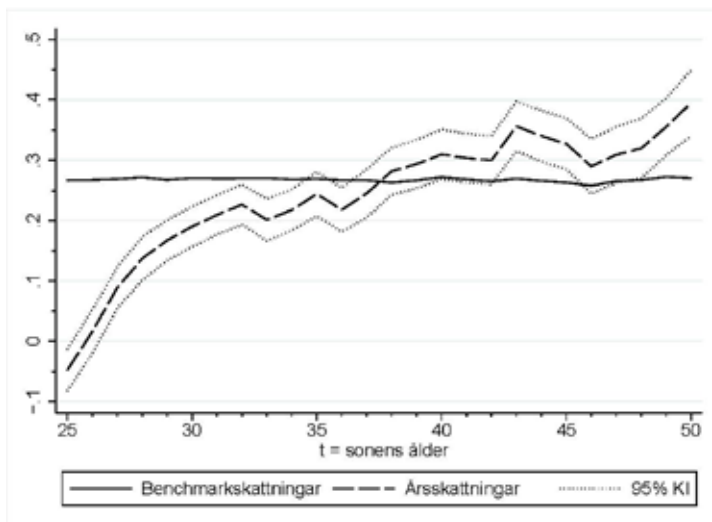
Slumpmässiga mätfel kan ge ett ”*bias mot noll*”, vilket innebär att det statistiska sambandet underskattas. Ett undantag gäller dock den intergenerationella elasticiteten, för vilken mätfel i söners inkomster (y^*) inte innebär något systematiskt ”*bias mot noll*”. Slumpmässiga mätfel påverkar dock de andra tre måtten och därför fortsätter de intergenerationella sambanden att vara underskattade även vid högre åldrar. Vi observerar dock att detta ”*bias mot noll*” är mycket mindre för rang- än för den linjära korrelationen. Rangkorrelationen är därför inte bara stabil över ålder utan även mycket närmare sitt benchmark från runt 30 års ålder och framåt. Den kvarvarande felskattning som vi ser är liten i jämförelse med den vid tidiga åldrar eller den vi ser för linjära mått. Den är dock inte helt obetydlig: runt 35–40 års ålder så uppgår den till 5–10 procent av benchmark-estimaten. I Figur 1 (d) tittar vi på sannolikheten att gå från den nedersta kvintilen (dvs. femtedelen) i föräldrafördelningen till den översta kvintilen i barnens inkomstfördelning, såsom i Chetty m.fl. (2014b, 2014a).³ Resultaten speglar de för rangkorrelationen: rörligheten är starkt överskattad innan 30 års ålder, men relativt välskattad och stabil därefter.

Dessa resultat illustrerar varför rangbaserade mått kan vara att föredra vid analys av intergenerationell inkomströrlighet i fall där mätproblem är viktiga att ta hänsyn till – vilket ofta är fallet. Vi fortsätter nu med en mer detaljerad diskussion av de fyra rörlighetsmåten.

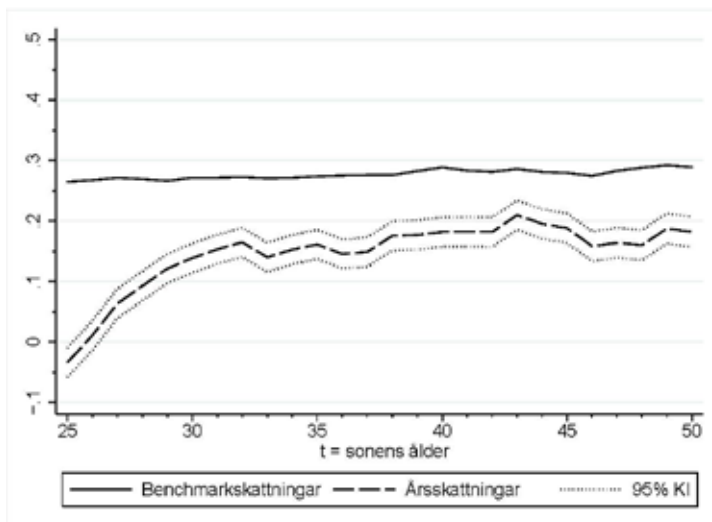
³ Dvs. sannolikheten att en son vars far tillhörde de fattigaste 20 procenten (bland fäder) själv kommer tillhöra de rikaste 20 procenten (i sin generation).

Figur 1 Rörlighetsmått, benchmark och ålderbaserade (mätfel i y^{*})

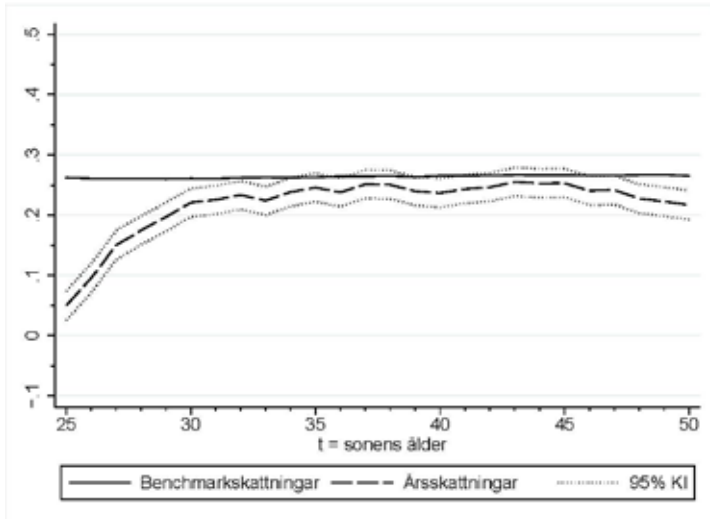
(a) Elasticiteten



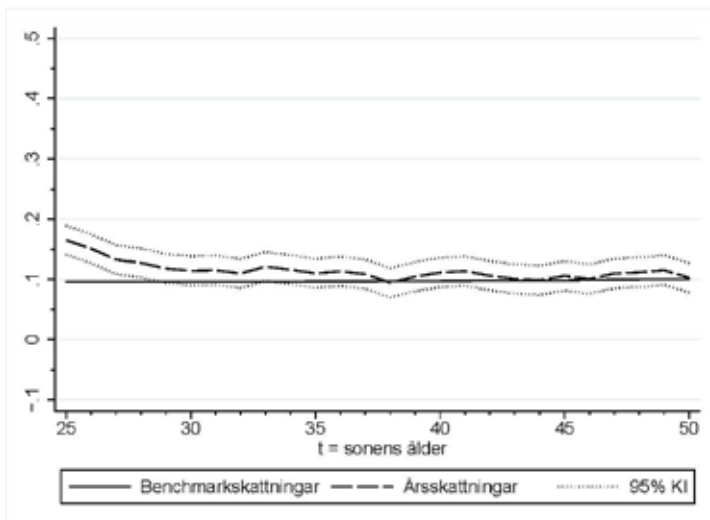
(b) Linjära korrelationen



(c) Rangkorrelationen



(d) Transitionssannolikhet, nedersta till översta femtedelen



Not: Respektive figur visar benchmark-estimat baserade på livsinkomster för båda generationer jämfört med estimat baserade på årsinkomster för söner. Stickprovet, och därför även benchmark-estimaten, varierar något över ålder då vi utesluter individer för vilka information om årsinkomst saknas även vid skattning av våra benchmark. Av visuella skäl visas inte konfidensintervall för benchmark-estimaten (de är lika eller något mindre i storlek jämfört med de åldersspecifika estimaten).

4 Fyra vanliga mått på intergenerationella samband i inkomster

4.1 Elasticiteten

Mycket av den deskriptiva rörlighetslitteraturen rör skattningen av den så kallade *intergenerationella elasticiteten*.⁴ Denna anger med hur många procent ett barns inkomst (y^*) förväntas avvika från genomsnittet i sin generations inkomstfördelning baserat på hur föräldrarnas inkomst (x^*) avviker från genomsnittet i föräldrarnas inkomstfördelning. Om elasticiteten till exempel är 0,25, vilket baserat på tidigare forskning är en rimlig uppskattning för Sverige, så innebär det att barn i *genomsnitt* ”ärver” ungefär 25 procent av sina föräldrars avvikelse från medelinkomsten. Ju högre denna elasticitet är, ju lägre är den intergenerationella rörligheten, och vice versa.

Den intergenerationella elasticiteten påverkas inte av slumpmässiga mätfel i söners inkomster. Då vi ovan tittade på mätfel för söner så är det bias som visas i Figur 1 därför endast en konsekvens av livscykeleffekter. Detta bias tenderar att öka med barnens ålder om de med höga livsinkomster har högre tillväxttakt i sina inkomster över tid. Men utvecklingen i inkomstspridning över ålder kan skilja sig åt mellan olika populationer, vilket gör jämförande analyser svåra.⁵

Det ”bias mot noll” som uppkommer som en konsekvens av mätfel i fäderns inkomster (x^*) illustreras i Figur 2. I denna figur jämförs, för vart och ett av de fyra olika rörlighetsmåten, benchmark-skattningar (heldragen linje) med skattningar baserade på söners årsinkomster vid olika åldrar *och* fäderns inkomster vid 45 års ålder (streckad linje), med eller utan korrigering för ”bias mot noll”. Benchmark-skattningarna visar de ”sanna” intergenerationella sambanden, och avvikelserna mellan dessa och de åldersspecifika skattningarna visar i vilken grad analyser baserade på årsinkomster för *både* fäder och söner leder till fel-skattningar. Korrigeringen som visas i figuren är standard i litteraturen och innebär att man justerar sina skattningar baserat på hur mycket slumpmässiga mätfel som den variabel man använder sig av innehåller.

4.2 Den linjära korrelationen

Den linjära (eller ”Pearson”) korrelationen skiljer sig åt från elasticiteten endast i det avseendet att den tar hänsyn till att inkomstspridningen kan förändras mellan generationer och korregerar för detta. På grund av denna korrigering får

⁴ Elasticiteten fås vanligtvis av lutningskoefficienten från en linjär regression av y^* på x^* .

⁵ Chetty m.fl. (2014a) rapporterar elasticitetsskattningar som stabiliseras runt 30 års ålder, medan Grawe (2006), Nybom & Stuhler (kommande), och Gregg m.fl. (2014) finner betydligt större åldersvariation.

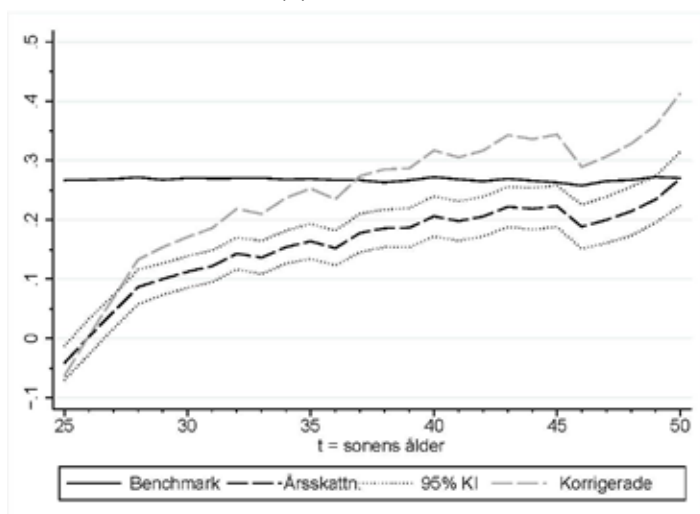
dock korrelationsskattningar ett ”bias mot noll” från slumpmässiga mätfel i *både* föräldrars och barns inkomster, medan det senare som sagt inte påverkade elasticiteten. Denna distinktion är av praktisk relevans; korrelationsskattningar är mycket lägre än elasticitetsskattningar vid mätfel i söners inkomster (Figur 1) men har ett liknande ”bias mot noll” vid mätfel i både faders och söners inkomster (Figur 2). Likt ovan kan man se detta genom att studera hur den streckade linjen skiljer sig åt från den heldragna linjen i respektive figur. Då den vanliga lösningen att använda flerårsgenomsnitt för föräldrars inkomster därför leder till stora förbättringar för elasticiteten, så är det för korrelationen lika viktigt att även ta hänsyn till slumpmässiga variationer i barnens inkomster.

Det tycks finnas ytterst lite, om ens någon, empirisk evidens kring s.k. livscykeffekter för intergenerationella korrelationer. Vi illustrerar i vårt working paper hur tidigare studier (t.ex. Haider & Solon 2006) av andra rörlighetsmått dock kan användas för att tolka våra resultat för korrelationen. Teoretiskt visar vi att korrelationsskattningar bör vara känsliga för livscykeffekter vid väldigt unga åldrar, men mindre känsliga i mitten och slutet av livscykeln.

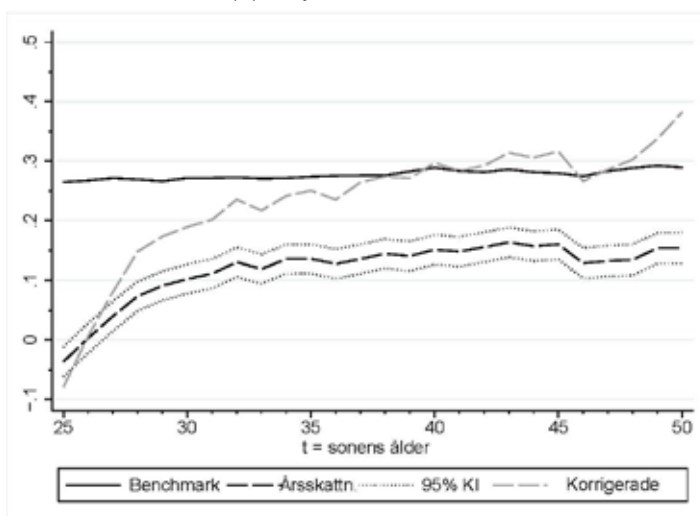
Våra empiriska resultat stödjer dessa hypoteser. Korrelationsskattningar har ett stort ”livscykelbias” när inkomster observeras vid unga åldrar, men håller sig relativt stabila efter trettioårsåldern. Detta gäller både vid mätfel i barn-generationens inkomster (Figur 1) och i båda generationernas inkomster (Figur 2). Även om korrelationsskattningar är mer stabila över ålder, så är nackdelen att de även vid senare åldrar underskattar graden av intergenerationell persistens (dvs. överskattar graden av rörlighet). Att korrigera för sådant ”bias mot noll” förbättrar skattningarna avsevärt, men leder i stället till mindre stabilitet över ålder. Detta framgår om den grå streckade linjen (de korrigerade skattningarna) jämförs med den svarta streckade linjen i Figur 2 (b). Vi ser då att de korrigerade skattningarna generellt ligger betydligt närmare benchmark-skattningarna (heldragen linje), men att de också uppvisar mer tydliga livscykeffekter; ju senare i livet man observerar söners inkomster, ju högre skattningar.

Figur 2 Rörlighetsmått, benchmark och ålderbaserade (mätfel i y^* och x^*)

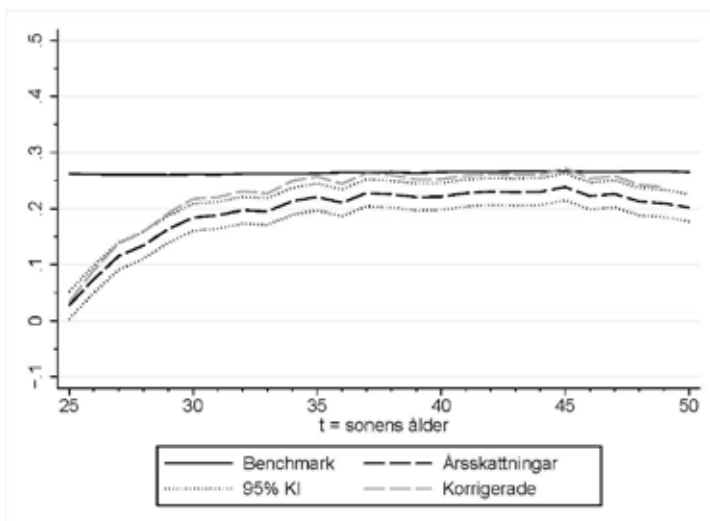
(a) Elasticiteten



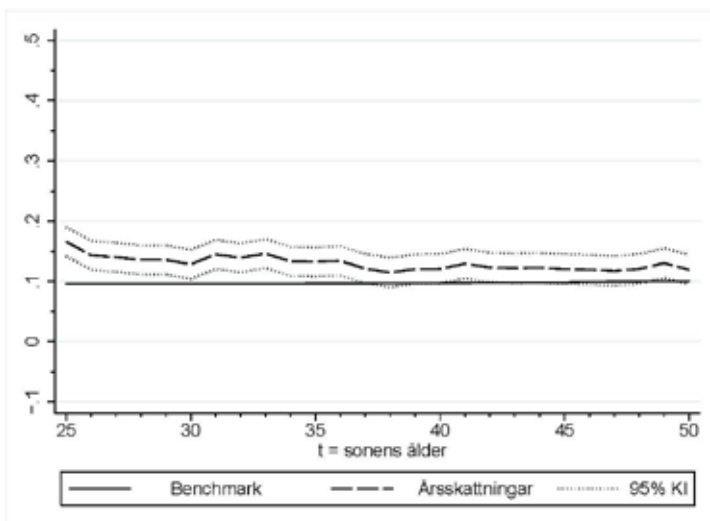
(b) Linjära korrelationen



(c) Rangkorrelationen



(d) Transitionssannolikhet, nedersta till översta femtedelen



Not: Respektive figur visar benchmark-estimat baserade på livsinkomster jämfört med estimat baserade på årsinkomster för söner (varierar med ålder) och fäder (vid 45 års ålder). Stickprovet, och därför även benchmark-estimatet, varierar något över ålder då vi utesluter individer för vilka information om årsinkomst saknas även vid skattning av våra benchmark. Figurerna (a)–(c) visar även estimat som justerats med hjälp av olika korrektioner för slumpmässiga mätfel (se avsnitt 4 och vårt working paper). Av visuella skäl visas inte konfidensintervall för benchmark-estimatet (de är lika eller något mindre i storlek jämfört med de åldersspecifika estimaten).

4.3 Rangkorrelationen

Rangbaserade mått fokuserar på individers position i inkomstfördelningen, ofta uttryckt som inkomstpercentiler. Till exempel har en person på 10:e inkomstpercentilen högre inkomst än 10 procent av befolkningen, medan en person på 50:e percentilen befinner sig mitt i inkomstfördelningen (dvs. vid medianen). I det aktuella fallet fokuserar vi på rangkorrelationen mellan söners och fäders inkomstpercentiler. Om rangkorrelationen är 0,50, till exempel, så betyder det att i genomsnitt hälften av faderns avvikelse från medianen ”ärvs” av sonen. En son till en rik far på 90:e percentilen förväntas alltså själv i genomsnitt hamna på 70:e percentilen, medan en son till en fattig far på den lägsta percentilen förväntas hamna nära 25:e percentilen. Om rangkorrelationen är noll förväntas samtliga söner att hamna på medianen (dvs. helt slumpmässigt); om den i stället är ett, då förväntas samtliga söner hamna på samma percentil som deras fäder. På samma vis som måtten ovan innebär alltså en låg rangkorrelation hög intergenerationell rörlighet, och vice versa.

Liksom den linjära korrelationen så hämmas även rangkorrelationen av ”bias mot noll” som en effekt av mätfel i båda generationers inkomster. Men är felskattningen större eller mindre? Vi diskuterar i detalj i det working paper denna rapport baseras på vad man rent teoretiskt kan förvänta sig. Sammanfattningsvis mynnar den diskussionen ut i att rangkorrelationen bör ha ett mindre ”bias mot noll” än den linjära korrelationen. Detta bekräftas även i Figur 1 och Figur 2, för vilka vi återigen kan analysera felskattningarna genom att jämföra de streckade linjernas avvikelse från de heldragna linjerna.

Rent matematiskt förväntar vi oss även att livscykeleffekterna är mindre för rangkorrelationen. Resultaten i Figur 1 och Figur 2 stödjer detta. Rangkorrelationen är, liksom de andra måtten, starkt underskattad vid unga åldrar. Detta då de med höga livsinkomster tenderar att ha relativt låga inkomster i början av deras karriärer. Rangkorrelationen är dock betydligt mindre påverkad av livscykeleffekter vid senare åldrar.

4.4 Transitionsmatriser och kopulan av barns och föräldrars inkomster

För att ytterligare illustrera fördelningen av föräldrars och barns inkomster så skattas ofta s.k. transitionsmatriser, vilka fångar rörligheten mellan olika delar av fördelningen (Zimmerman, 1992; Jäntti m.fl., 2006). Fokus ligger här ofta på specifika delar av fördelningen, till exempel den intergenerationella persistensen (eller ”örörligheten”) i den absoluta toppen eller botten av fördelningen, eller på sannolikheten att stiga från botten till toppen. Som exempel kan man här analysera sannolikheten att tillhöra de rikaste 10 procenten givet

att ens förälder tillhörde de rikaste 5 eller 10 procenten ("ärvandet av toppinkomster"), sannolikheten att förbli fattig givet att ens förälder var fattig ("fattigdomsfällor"), eller sannolikheten att nå högt i inkomstfördelningen trots att ens förälder var fattig ("klassresor"). Enkelt uttryckt så består en transitionsmatris av en lång rad av sådana sannolikheter, vilka definieras utifrån hur man väljer att dela upp inkomstfördelningen.

Med tillräckligt stora datamaterial kan vi urskilja många olika inkomstklasser och approximera den s.k. kopulan, den gemensamma fördelningen av föräldrars och barns rangordningar (dvs. inkomstpercentiler). Kopulan visar sannolikheten att sonen återfinns i olika delar av inkomstfördelningen (vanligtvis i termer av inkomstpercentiler mellan 1 och 100), givet faderns motsvarande position i inkomstfördelningen bland fäder. Då kopulan är en underliggande bestämmelsefaktor av samtliga rörlighetsmått (se Chetty m.fl. 2014b, 2014a) så är vetenskapen om hur välskattad den är med mer temporära inkomstmått (såsom årsinkomster) värdefull.

Vi diskuterar skattning av transitionsmatriser och kopulan i mer detalj i vårt working paper. I korthet ger vår teoretiska diskussion att den robusthet vi tidigare fann för rangkorrelationen inte nödvändigtvis håller för hela fördelningen. Vi förväntar oss framför allt att rörligheten i botten och toppen, samt sannolikheten för långväga förflyttningar (såsom från botten av föräldrafördelningen till toppen av barnens fördelning), kan överskattas när kortsiktiga inkomstdata används.

Vi undersöker detta genom att skatta kopulan med våra data. Dock är sådan analys betydligt mer dataintensiv än de som presenterats ovan. Vårt datamaterial kan utökas betydligt om vi mäter föräldrars inkomster från 1968 i stället för 1960, så att faders inkomster observeras endast från 41 års ålder och framåt (se avsnitt 2).⁶

Figur 3 visar kopulan av föräldrars och barns långsiktiga inkomster. Figuren är i princip en illustration av ett tredimensionellt förhållande, där faders och söners inkomstfördelningar ses på de två axlarna och färgerna visar sannolikheten för en viss kombination (ju mörkare färg, ju högre sannolikhet). Sannolikheten är högst längs diagonalen: inte oväntat så är den mest sannolika positionen i inkomstfördelningen samma för barn som för deras föräldrar. Då rangordnade variabler har en högsta och lägsta gräns så är denna sannolikhet särskilt hög för barn med de allra fattigaste och (framför allt) rikaste föräldrarna. Sannolikheten att nå den högsta inkomstpercentilen om även fadern tillhörde denna percentil är 12,5 procent i våra data; samma grupps chans att nå någon av de högsta fem

⁶ I våra data är korrelationen mellan rangordningar baserade på "fullständiga" livsinkomster (åldern 25–65) och kortare mått 0,96 (åldern 31–60) respektive 0,91 (åldern 41–60).

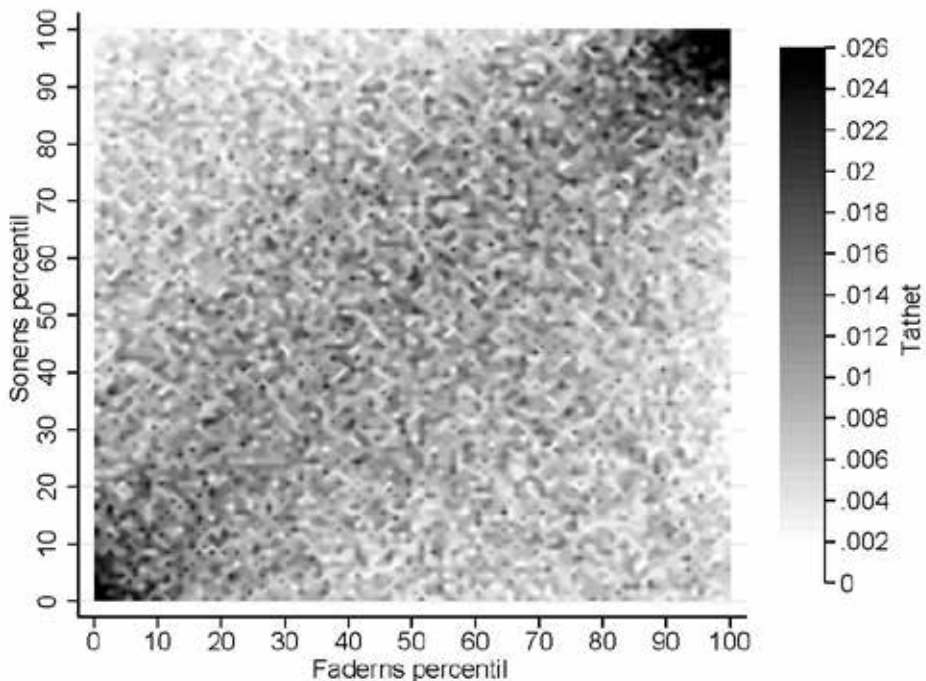
percentilerna är över 38 procent. För att få en meningsfull bild av övriga delar av fördelningen visas inte den högsta percentilen i Figur 3.

I vårt working paper undersöker vi även för kopulan de felskattningar som uppstår när man använder årsinkomster för söner, eller för både söner och fäder. Här sammanfattas dessa resultat endast kortfattat. För det första, användandet av årsinkomster leder till underskattning längs diagonalen och överskattning utanför diagonalen – i linje med vad vi förväntat oss teoretiskt. För det andra, felskattningarna är störst i toppen och botten av fördelningen. Medan ärvandet av fattigdom respektive toppinkomster ofta är av särskilt intresse, så är alltså just sådana skattningar också minst pålitliga med korta inkomstdata.

Våra resultat är således i linje med simuleringsresultaten i O'Neill m.fl. (2007), vilka antyder att slumpmässiga mätfel framför allt leder till överskattning av rörligheten i toppen och botten av fördelningen. Även graden av långväga rörlighet nedåt i fördelningen, t.ex. från toppen till botten, visar sig vara relativt känslig. Då även söner till väldigt rika föräldrar vid enstaka tillfällen kan ha låga årsinkomster så är sannolikheten att dessa faller till botten starkt överskattad i data baserade på årsinkomster.⁷ Vi anser därför att effekten av mätfel på kopulan har en ganska logisk struktur och att man bör vara särskilt varsam när man studerar långväga rörlighet och persistensen i botten och toppen av inkomstfördelningen.

⁷ Detta problem är troligtvis större när inkomster observeras över ännu kortare tidsperioder (t.ex. vecko- eller månadsvis), vilket är vanligt i surveybaserade datakällor.

Figur 3 Sannolikhetsfördelningen av fäders och söners inkomster (benchmark)



Not: Figuren visar kopulan, dvs. sannolikhetsfördelningen av söners och fäders inkomst-percentiler, baserat på livsinkomster för båda generationer. Under statistiskt oberoende så har varje cell sannolikheten (eller "tätheten") 0,01. Ju mörkare färg en cell har, ju högre sannolikhet (eller "täthet"). Sannolikheter längs diagonalen fångar "örörlighet" (dvs. persistens) och sannolikheter utanför diagonalen fångar rörlighet.

5 Slutsatser

Vi undersöker olika typer av felskattningar i fyra vanligt förekommande mått på intergenerationell rörlighet i inkomster med hjälp av närapå heltäckande livsinkomstdata för svenska fäder och deras söner. Vi finner att persistensen (rörligheten) är gravt underskattad (överskattad) för alla mått om kortsiktiga inkomster, såsom årsinkomster, från innan 30 års ålder används. Dock varierar graden av och riktningen på felskattningarna mycket vid senare åldrar. Elasticitetsskattningar påverkas mycket av s.k. livscykeeffekter, vilket innebär att de överskattar rörligheten när barnens inkomster mäts tidigt i yrkeslivet, och underskattar rörligheten när de mäts sent. Vanliga linjära korrelationer är mer stabila över ålder men också konstant gravt underskattade.

Rangbaserade mått såsom rangkorrelationen och diverse transitionssannolikheter är felskattade i betydligt mindre utsträckning. Särskilt positivt är den stabilitet över ålder som dessa rangbaserade mått uppvisar så länge inkomster observeras efter 30 års ålder. Även om viss underskattning kvarstår så är den betydligt mindre i storlek än för de mer traditionella måtten (elasticiteten och linjära korrelationen). Dock visar vi att även rangbaserade mått kan vara relativt opålitliga i toppen och botten av inkomstfördelningen. Rörligheten är slående låg både bland de allra rikaste och allra fattigaste i våra data för Sverige. I mitten av fördelningen är den betydligt högre. Dessa typer av skillnader i olika delar av fördelningen är viktiga; den absoluta toppen innehar en oproportionerligt stor del av totalinkomsten i de flesta länder och vi kan av flera anledningar bry oss extra mycket om rörligheten bland de fattigaste, för vilka även små inkomstförändringar kan ha stora välfärdseffekter. Hög persistens i botten och toppen kan också relatera till specifika mekanismer, såsom ärvandet av företag, kapital och arbetsgivare eller, å andra sidan, kreditrestriktioner eller ärvandet av långtidsarbetslöshet, dåliga kunskaper och låg utbildning, eller hälsoproblem. Linjära mått fångar inte sådana mönster. Våra resultat antyder dock att även analyser med fokus på olika delar av fördelningen kan missa att fånga dessa mönster om analyserna baseras på kortsiktiga inkomstmått (i vårt fall årsinkomster).

Hur robusta olika rörlighetsmått är kan dock bara ses som en faktor av flera. Den konceptuella relevansen av olika mått behöver också tas hänsyn till. Den intergenerationella elasticiteten brukar till exempel av många ses som mer intuitiv då den kan härledas från standardmodeller rörande hur föräldrar investerar i sina barn (Becker & Tomes 1979, Solon 2004). Rangbaserade mått kan istället ha ganska olika betydelse i olika kontexter och för olika inkomstfördelningar. Att flytta sig tio percentiler i ett land som USA, med stora inkomstskillnader, har andra implikationer än motsvarande förflyttning i ett land som Sverige. Vi hoppas dock att våra resultat kan vara ett stöd när mätproblematik och sådana konceptuella hänsynstaganden ställs mot varandra. Åtminstone bör våra resultat ge en anledning till att använda rangbaserade mått som komplement till eller känslighetsanalyser av de mer traditionella måtten.

Referenser

- Becker, G., & Tomes, N. (1979). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy* , 87(6), s. 1153–1189.
- Björklund, A., & Jäntti, M. (2009). Intergenerational Income Mobility and the Role of Family Background. I B. N. Salverda Wiemer, & T. M. Smeeding (red.), *Oxford Handbook of Economic Inequality*. Oxford University Press.
- Black, S. E., & Devereux, P. (2011). Recent Developments in Intergenerational Mobility. I O. Ashenfelter, & D. Card (red.), *Handbook of Labor Economics* (Vol. 4A). Elsevier.
- Bratberg, E., Davis, J., Mazumder, B., Nybom, M., Schnitzlein, D., & Vaage, K. (2015). A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the U.S. Discussion Paper, mimeo.
- Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014a). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *Quarterly Journal of Economics*, 129(4), s. 1553–1623.
- Chetty, R., Hendren, N., Kline, P., Saez, E., & Turner, N. (2014b). Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility. *American Economic Review: Papers and Proceedings*, 104(5), s. 141–147.
- Corak, M., Lindquist, M. J., & Mazumder, B. (2014). A Comparison of Upward and Downward Intergenerational Mobility in Canada, Sweden and the United States. *Labour Economics*.
- Grawe, N. D. (2006). Lifecycle Bias in Estimates of Intergenerational Earnings Persistence. *Labour Economics*, 13(5), 551–570.
- Gregg, P., MacMillan, L., & Vittori, C. (2014). Moving Towards Estimating Lifetime Intergenerational Economic Mobility in the UK. DoQSS Working Papers 14-12, Institute of Education, University of London.
- Haider, S., & Solon, G. (2006). Life-Cycle Variation in the Association Between Current and Lifetime Earnings. *American Economic Review* , 96(4), s. 1308–1320.
- Jenkins, S. (1987). Snapshots versus Movies: Lifecycle Biases' and the Estimation of Intergenerational Earnings Inheritance. *European Economic Review* , 31(5), s. 1149–1158.

- Jäntti, M., & Jenkins, S. P. (2014). Income Mobility. I A. B. Atkinson, & F. Bourguignon (red.), *Handbook of Income Distribution* (Vol. 2). Elsevier.
- Jäntti, M., Bratsberg, B., Roed, K., Naylor, R., Osterbacka, E., Björklund, A., & Eriksson, T. (2006). American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States. IZA Discussion Papers 1938.
- Nybom, M., & Stuhler, J. (forthcoming). Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation. *Journal of Human Resources*.
- Nybom, M., & Stuhler, J. (2015). Biases in Standard Measures of Intergenerational Income Dependence. IFAU Working paper 2015:13.
- O'Neill, D., Sweetman, O., & Gaer, D. v. (2007). The Effects of Measurement Error and Omitted Variables when Using Transition Matrices to Measure Intergenerational Mobility. *Journal of Economic Inequality*, 5(2), s. 159–178.
- Pekkarinen, T., Salvanes, K. G., & Sarvimäki, M. (2015). The Evolution of Social Mobility: Norway over the 20th Century. Discussion Paper, mimeo.
- Solon, G. (2004). A Model of Intergenerational Mobility Variation over Time and Place. I M. Corak (red.), *Generational Income Mobility in North America and Europe* (s. 38–47). Cambridge University Press.
- Solon, G. (1999). Intergenerational Mobility in the Labor Market. I O. Ashenfelter, & D. Card (red.), *Handbook of Labor Economics* (Vol. 3A, s. 1761–1800). Elsevier.
- Zimmerman, D. J. (1992). Regression Toward Mediocrity in Economic Stature. *American Economic Review*, 82(3), s. 409–429.

IFAU:s publikationsserier – senast utgivna

Rapporter

- 2015:1** Albrecht James, Peter Skogman Thoursie och Susan Vroman ”Glastaket och föräldraförsäkringen i Sverige”
- 2015:2** Persson Petra ”Socialförsäkringar och äktenskapsbeslut”
- 2015:3** Frostenson Magnus ”Organisatoriska åtgärder på skolnivå till följd av lärarlegitimationsreformen”
- 2015:4** Grönqvist Erik och Erik Lindqvist ”Kan man lära sig ledarskap? Befälsutbildning under värnplikten och utfall på arbetsmarknaden”
- 2015:5** Böhlmark Anders, Helena Holmlund och Mikael Lindahl ”Skolsegregation och skolval”
- 2015:6** Håkanson Christina, Erik Lindqvist och Jonas Vlachos ”Sortering av arbetskraftens förmågor i Sverige 1986–2008”
- 2015:7** Wahlström Ninni och Daniel Sundberg ”En teoribaserad utvärdering av läroplanen Lgr 11
- 2015:8** Björvang Carl och Katarina Galic’ ”Kommunernas styrning av skolan – skolplaner under 20 år”
- 2015:9** Nybom Martin och Jan Stuhler ”Att skatta intergenerationella inkomst-samband: en jämförelse av de vanligaste måtten”

Working papers

- 2015:1** Avdic Daniel “A matter of life and death? Hospital distance and quality of care: evidence from emergency hospital closures and myocardial infarctions”
- 2015:2** Eliason Marcus “Alcohol-related morbidity and mortality following involuntary job loss”
- 2015:3** Pingel Ronnie och Ingeborg Waernbaum “Correlation and efficiency of propensity score-based estimators for average causal effects”
- 2015:4** Albrecht James, Peter Skogman Thoursie och Susan Vroman “Parental leave and the glass ceiling in Sweden”
- 2015:5** Vikström Johan “Evaluation of sequences of treatments with application to active labor market policies”
- 2015:6** Persson Petra “Social insurance and the marriage market”
- 2015:7** Grönqvist Erik och Erik Lindqvist “The making of a manager: evidence from military officer training”
- 2015:8** Böhlmark Anders, Helena Holmlund och Mikael Lindahl “School choice and segregation: evidence from Sweden”

- 2015:9** Håkanson Christina, Erik Lindqvist och Jonas Vlachos “Firms and skills: the evolution of worker sorting”
- 2015:10** van den Berg Gerard J., Antoine Bozio och Mónica Costa Dias “Policy discontinuity and duration outcomes”
- 2015:11** Wahlström Ninni och Daniel Sundberg “Theory-based evaluation of the curriculum Lgr 11”
- 2015:12** Frölich Markus och Martin Huber “Direct and indirect treatment effects: causal chains and mediation analysis with instrumental variables”
- 2015:13** Nybom Martin och Jan Stuhler “Biases in standard measures of inter-generational income dependence”

Dissertation series

- 2014:1** Avdic Daniel “Microeconomic analyses of individual behaviour in public welfare systems”
- 2014:2** Karimi Arizo “Impacts of policies, peers and parenthood on labor market outcomes”
- 2014:3** Eliasson Tove “Empirical essays on wage setting and immigrant labor market opportunities”
- 2014:4** Nilsson Martin “Essays on health shocks and social insurance”
- 2014:5** Pingel Ronnie “Some aspects of propensity score-based estimators for causal inference”
- 2014:6** Karbownik Krzysztof “Essays in education and family economics”