

Bilaga 5.

Kvantitativ studie



RiR 2020:13

Stöd och matchning

– ett valfrihetssystem för arbetssökande

Innehåll

1	Data	3
2	Leverantörerna	4
2.1	Utvecklingen över tid	5
2.2	Leverantörer och leveransområden	6
2.3	Sambandet mellan antal arbetslösa och antal leverantörer	8
3	Deltagarna	11
3.1	Deltagarsammansättningen över tid	11
3.2	Beskrivning av deltagarna uppdelat på rating	13
4	Ratingen	15
5	Incitament för effektivt beteende hos leverantörer	19
5.1	Datamaterialet och ratingen	19
5.2	Lönar det sig med en hög rating?	22
5.3	Samvariationen mellan rating och leverantörens storlek och tid i systemet	34
5.4	Kvalitet i R2 mellan leverantörerna	35
	Referenser	39

I denna bilaga genomförs ett antal kvantitativa analyser av STOM och dess delar. Vi börjar med att beskriva dataunderlaget, följt av ett deskriptivt fokus på leverantörerna där även en enklare korrelationsanalys genomförs på leveransområdesnivå. Sedan följer ett deskriptivt avsnitt om deltagarna i STOM. Efterföljande kapitel beskriver ratingmodellen och hur den tas fram. Det kapitlet tjänar som bakgrund till det mer analytiska kapitlet om hur incitamenten ser ut i STOM-systemet. Mer specifikt analyseras hur ratingen påverkar systemöverlevnad och deltagarinflöde för leverantörerna. Dessutom undersöks kvaliteten i det resultatmätt som Arbetsförmedlingen använder för att resultatersättning ska utbetalas till leverantören, det så kallade R2. Detta undersöks genom att jämföra hur tiden i inrapporterad anställning korresponderar med Skatteverkets månadsuppgifter samt hur godkänd R2 och den registerbaserade sysselsättningen på ett års sikt skiljer sig mellan individer som fått R2 genom leverantörer med olika rating.

1 Data

Dataunderlaget utgörs dels av Arbetsförmedlingens offentligt tillgängliga aggregerade data från deras webbplats, dels av mikrodata beställt från Arbetsförmedlingen, Statistiska centralbyrån och Skatteverket.

Mikrodata från Arbetsförmedlingen (AF) avser alla leverantör med giltiga avtal och minst ett beslut hos AF mellan december 2014 och september 2019 (uppgifter om rating finns dock först från juni 2016). Till detta tillkommer även individuppgifter om deltagarna hos dessa leverantörer som fick godkänt R2 (det vill säga som funnit arbete eller påbörjat studier). Till dessa individer har data från SCB (LISA) och Skatteverkets månadsuppgifter kopplats.

2 Leverantörerna

Under 12-månadersperioden mellan december 2018 och november 2019 fanns det 188 unika leverantörer med minst 1 deltagare under året. Tabell A1 visar beskrivande statistik över dessa leverantörer. Här framgår att medelantalet deltagare var 77 och medianen var 33. Fördelningen är alltså snedfördelad till höger, vilket innebär att det finns ett fåtal leverantörer som är betydligt större än medianleverantören. Spridningen över landet är ungefär lika snedfördelad, variationskoefficienten (standardavvikelse/medelvärde) är 1,72 för deltagarna och 1,69 för leveransområdena.

Tabell A1 Deskription över leverantörer (dec 2018–nov 2019)

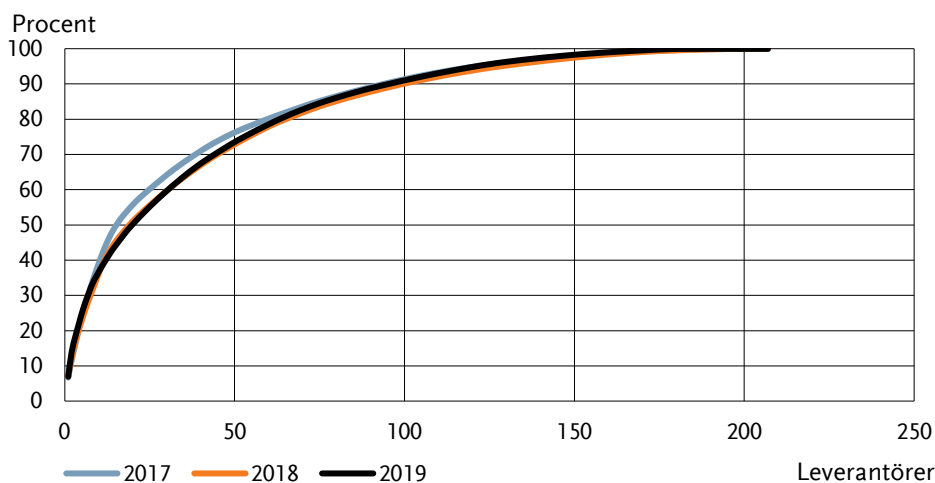
Deltagare per månad					Unika leverantörer	Leveransområde		
Medelvärde	SD	p10	Median	p90	Antal	Medelvärde	SD	Median
77	132	7	33	153	188	3,5	6	2

Anm.: SD = Standardavvikelse, p10 = 10:e percentilen, p90 = 90:e percentilen.

Källa: Arbetsförmedlingen.

De 10 största unika leverantörerna hade i genomsnitt 37 procent av alla deltagare per månad under 2019 (se figur A1). Detta förhållande har varit relativt stabilt över tid.

Figur A1 Kumulativ andel deltagare per månad för leverantörer fördelat på år, rangordnat efter deltagarantal

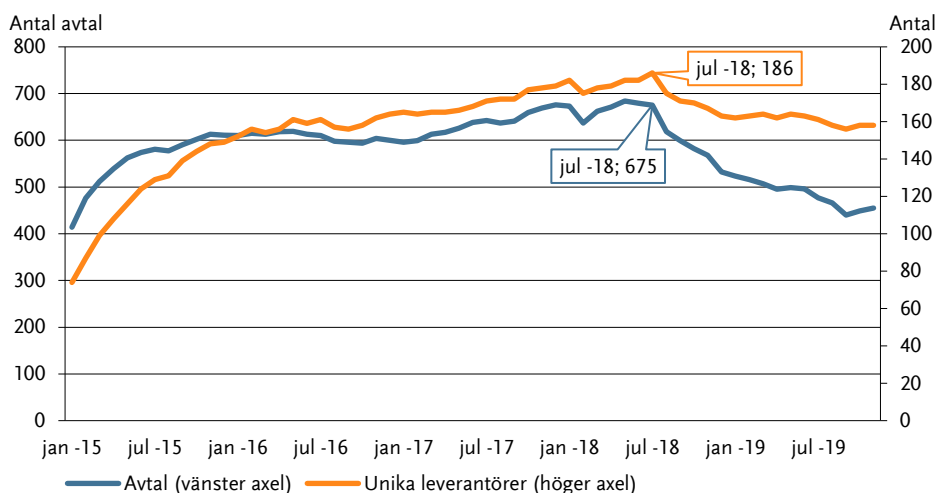


Källa: Arbetsförmedlingens webbplats.

2.1 Utvecklingen över tid

I november 2019 fanns det 158 unika leverantörer (organisationsnummer) i STOM fördelat på 58 av 72 möjliga leveransområden (FA05). Eftersom en organisation kan finnas på flera olika leveransområden blir det sammantaget 455 aktiva avtal¹ i systemet samma månad. Systemet var betydligt större tidigare men från juli 2018 fram tills idag har antalet aktiva leverantörer och avtal minskat tydligt.

Figur A2 Antal leverantörer och avtal i systemet

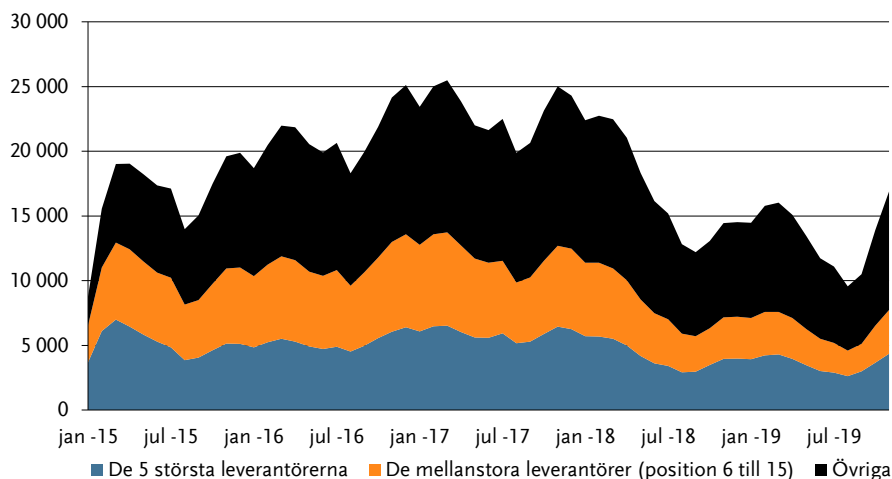


Källa: Arbetsförmedlingens webbplats.

Figur A3 visar att även antalet deltagare i STOM började minska redan våren 2018. En förklaring till denna minskning var det minskade anslaget till Arbetsförmedlingen för 2018 som ledde till att Arbetsförmedlingen prioriterade resurserna till arbetssökande inom jobb- och utvecklingsgarantin, jobbgarantin för ungdomar och etableringsprogrammet.² Denna minskning skedde främst inom de största leverantörerna. Mellan mars och juli 2018 minskade antalet deltagare med 38, 34 och 29 procent för de största, mellanstora och övriga leverantörerna. Men sett på lite längre sikt är det mellansegmentet som förlorat flest deltagare. Den största unika leverantören innan våren 2018 hade omkring 1 700 deltagare per månad medan den största leverantören idag har cirka 1 300 deltagare.

¹ Med aktiva avtal avses de avtal som har fått något beslut någon gång, dvs. minst en deltagare någon gång. Avtal och KA-nummer är samma sak (bortsett från ett fåtal felregistreringar). Leverantörer som är aktiva i olika leveransområden får olika KA-nummer i de olika leveransområdena.

² Uppgifter från Arbetsförmedlingen i e-post 2020-01-10 och 2019-12-13.

Figur A3 Deltagare fördelat på leverantörers storlek

Anm.: En leverantörs storlek baseras på antalet deltagare. Även om en leverantör är verksam i flera leveransområden räknas de här som en enda leverantör.

Källa: Arbetsförmedlingens webbplats.

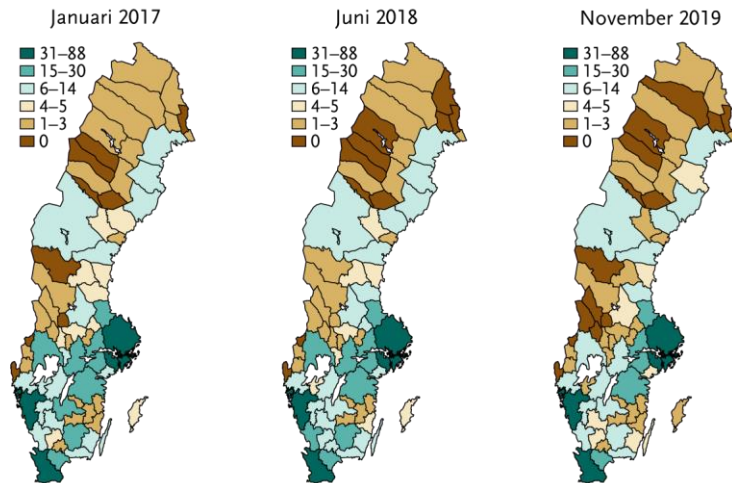
2.2 Leverantörer och leveransområden

Arbetsförmedlingen tecknar avtal med leverantörerna per leveransområde. Totalt är landet uppdelat i 72 leveransområden, som är baserade på Tillväxtverkets så kallade FA-regioner från 2005.³ En leverantör kan vara verksam i ett eller flera leveransområden.

De kraftigt minskade volymerna i systemet (både leverantörer och deltagare) har lett till att allt fler leveransområden har få eller helt saknar leverantörer (se figur A4).

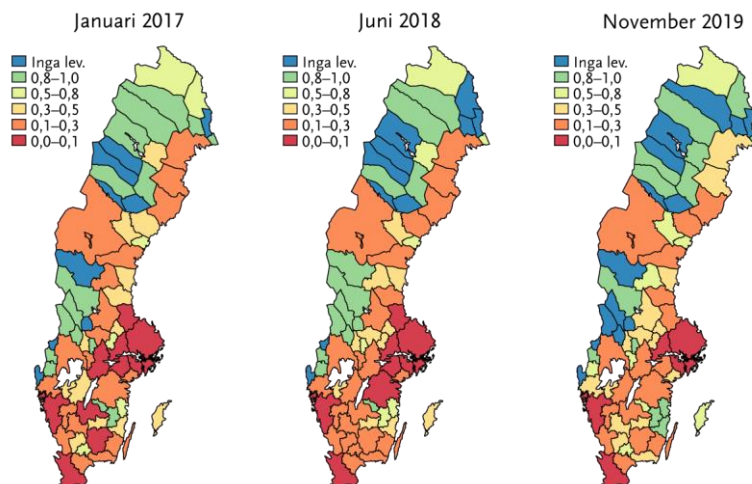
I november 2019 fanns det 42 (58 procent) leveransområden med högst 3 leverantörer och 14 (19 procent) utan någon leverantör alls. Dessa områden omfattar 7,4 respektive 1,1 procent av befolkningen mellan 20 och 64 år (beräknat för 2018). Systemets täckningsgrad är med andra ord betydligt bättre på befolkningsnivå än vad en beräkning på den geografiska ytan visar.

³ Arbetsförmedlingen, *Kundval - Stöd och matchning*, 2018.

Figur A4 Antal leverantörer per leveransområde

Källa: Arbetsförmedlingens webbplats.

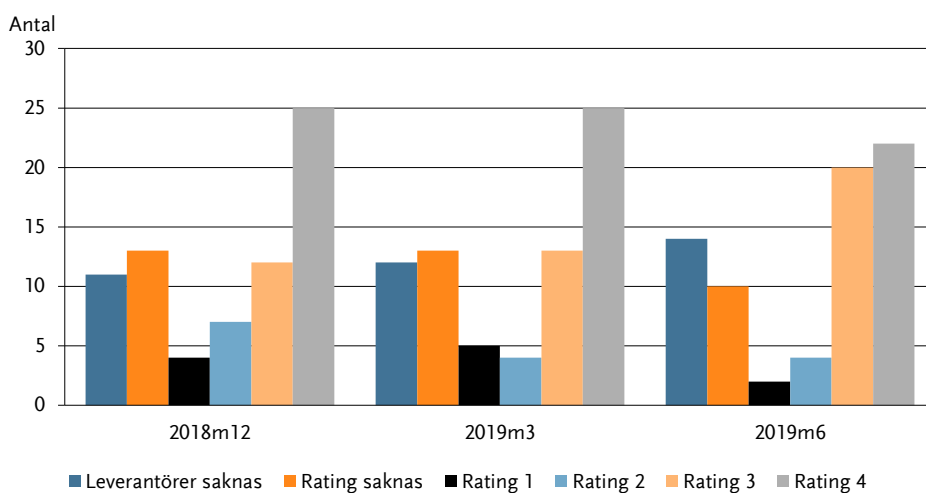
Herfindahl index är en vanlig metod för att beräkna marknadskoncentration. Indexet går mellan 0 och 1, där ett tal nära 0 innebär en låg marknadskoncentration och därmed hög konkurrens medan ett tal nära 1 innebär att det i princip finns en enda aktör på den marknaden. Figur A5 visar hur marknadskoncentrationen har förändrats över tid. I genomsnitt har marknadskoncentrationen ökat (om vi exkluderar de regioner där leverantörer helt saknas under dessa mättidpunkter). Januari 2017 låg indexet på cirka 0,38 medan den november 2019 låg på cirka 0,44.

Figur A5 Lokal marknadsdominans (index)

Källor: Arbetsförmedlingens webbplats och egna beräkningar.

I stora delar av landet är konkurrensen låg eller saknas helt vilket. I dessa områden fungerar således inte de marknadsmekanismer som ratingsystemet⁴ var menat att stärka. För det senaste fullständiga estimeringskvartalet (juni till augusti 2019) fanns det 12 leveransområden som helt saknade leverantörer och 13 områden med leverantörer men där alla saknade en rating (se figur A6). I 5 respektive 4 områden hade ingen leverantör mer än 1 eller 2 i rating. I 13 områden fanns ingen leverantör med mer än 3 stjärnor. I 25 områden hade deltagarna möjlighet att välja en topprankad⁵ leverantör med 4 stjärnor. Cirka 74 procent av alla nya STOM-deltagare fanns i dessa områden med minst en topprankad leverantör.

Figur A6 Antal områden där det finns minst en leverantör med ingen 1, 2, 3 eller 4 i rating



Anm.: Notera att en rating beräknas per avtal, dvs. per leverantör och leveransområde.

Källa: Arbetsförmedlingens webbplats.

2.3 Sambandet mellan antal arbetslösa och antal leverantörer

För att undersöka leverantörers etablering inom ett specifikt leveransområde beräknar vi en enkel regressionsmodell där antalet leverantörer förklaras av hur många arbetssökande det finns per kvadratkilometer i ett leveransområde. Denna variabel används eftersom det är rimligt att tänka sig att det krävs ett visst antal potentiella deltagare i form av arbetslösa individer för att en leverantör ska etablera sig i ett område.⁶ Därefter görs ett räkneexempel för att undersöka vad som skulle krävas för att få tre leverantörer i samtliga leveransområden. Vi har valt tre leverantörer för att det kan ses som ett lägsta krav för vad vi skulle kunna beteckna som konkurrens.

⁴ För en beskrivning av ratingen, se sida 14.

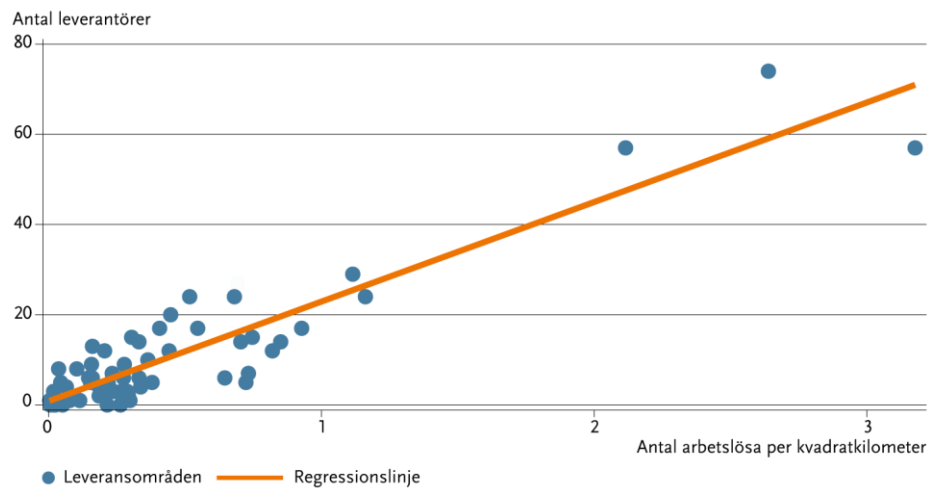
⁵ I vissa fall lämpar sig inte olika ordformer av *rating*. I dessa fall kommer böjningar av synonymen *ranking* att användas i stället. Exempel: "de högst ~~rate~~ rankade leverantörerna".

⁶ Arbetssökande per kvadratkilometer är således vår valda proxy för potentiellt antal STOM-deltagare. En proxy är definitionsmässigt inte ett exakt mått på vad vi egentligen vill mäta, exempelvis är inte alla arbetslösa egentligen potentiella deltagare för STOM. Den viktigaste bestämningsfaktorn för antalet potentiella deltagare är dock den arbetsföra befolkningen (mellan exempelvis 20 och 64 år), den lokala arbetsmarknadsläget och sammansättningen av de arbetslösa spelar en mindre roll i sammanhanget.

Antalet arbetsökande är hämtade från SCB:s databas LISA och avser situationen i november 2017 och jämförs därför med leverantörsetableringen inom STOM för samma månad. I den studerade perioden var STOM betydligt större än det är idag (fler leverantörer och deltagare). Dessa beräkningar är alltså utförda i ett relativt stabilt och gynnsamt läge för systemet.

I figur A7 presenteras sambandet mellan antal arbetslösa per kvadratkilometer inom ett leveransområde och antalet etablerade leverantörer inom samma leveransområde. Det framkommer att antal arbetslösa per kvadratkilometer tycks kunna predicera hur många leverantörer som finns i ett leveransområde relativt väl. Den mycket enkla regressionsmodellen kan förklara knappt 85 procent av variationen i antal leverantörer.

Figur A7 Sambandet mellan antal arbetslösa per kvm och antalet etablerade leverantörer



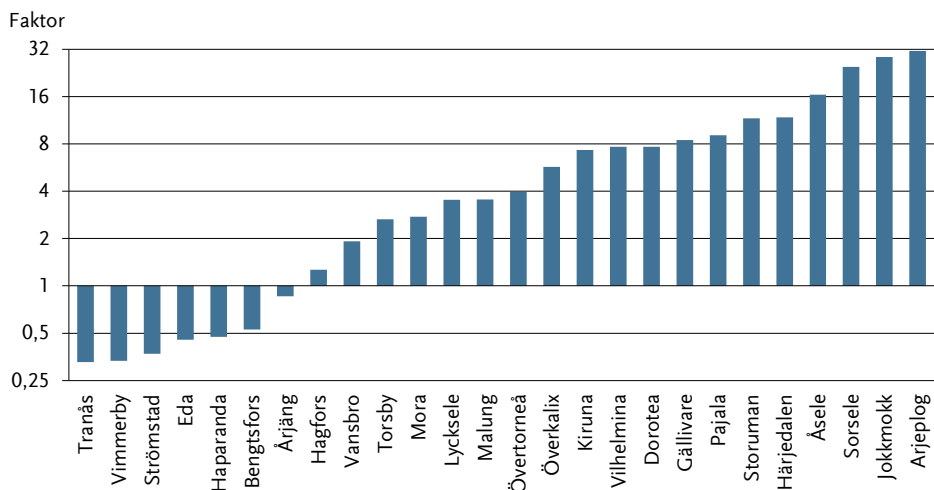
Källor: Arbetsförmedlingen, Statistiska centralbyrån och egna beräkningar.

Utifrån resultaten från denna modell görs sedan ett räkneexempel för att illustrera hur många fler arbetsökande per kvadratkilometer som, enligt modellen, skulle behövas för att samtliga leveransområden ska få minst 3 etablerade leverantörer. Sammantaget var det 26 leveransområden som hade färre än 3 leverantörer i november 2017. Samtidigt fanns det ingen leverantör i 9 leveransområden. Resultaten från räkneexemplet presenteras i figur A8.

Det skulle enligt den beräknade modellen krävas en stor ökning av antalet arbetslösa per kvadratkilometer. För att nå minst 3 leverantörer i alla dessa leveransområden behöver antalet arbetslösa per kvadratkilometer öka 7 gånger om. I exempelvis Arjeplog skulle enligt räkneexemplet antalet arbetslösa per kvadratkilometer behöva vara drygt 30 gånger högre för att få 3 leverantörer att etablera sig. Samtidigt kan konstateras att det enligt exemplet finns ett antal områden som borde haft 3 leverantörer redan med det existerande underlaget av arbetsökande per kvadratkilometer som fanns i november 2017. Detta gäller exempelvis Tranås, Vimmerby och Strömstad.

Sammantaget tycks det enligt detta enkla räkneexempel krävas ganska många fler potentiella deltagare för att få till en konkurrenssituation med tre leverantörer i samtliga områden. Siffrorna i räkneexemplet kan även användas för att bilda sig en uppfattning om hur mycket andra ingående parametrar i den underliggande etableringsekvationen⁷ som en leverantör möter måste förändras för att nå åtminstone 3 etablerade leverantörer per område. Eftersom kundunderlaget i exempelvis Dorotea är givet på förhand är det snarare möjligt att minska exempelvis fasta och rörliga kostnader knutna till de krav på leverantörerna som Arbetsförmedlingen satt upp. Detta kan givetvis även kombineras med en ökad betalning per deltagare eller någon sorts fast etableringspremie satt till respektive område.

Figur A8 Förändring av antalet arbetslösa per kvadratmeter som skulle krävas för nå tre etablerade leverantörer i samtliga leveransområden



Anm.: Y-axeln är i logaritmisk skala med basen 2 för att underlätta läsningen och synliggöra de områden vars värden är mindre än 1. Ett värde mindre än 1 innebär att dessa områden, enligt modellen, redan har förutsättningar för minst 3 leverantörer i området.

Källor: Arbetsförmedlingen, Statistiska centralbyrån och egna beräkningar.

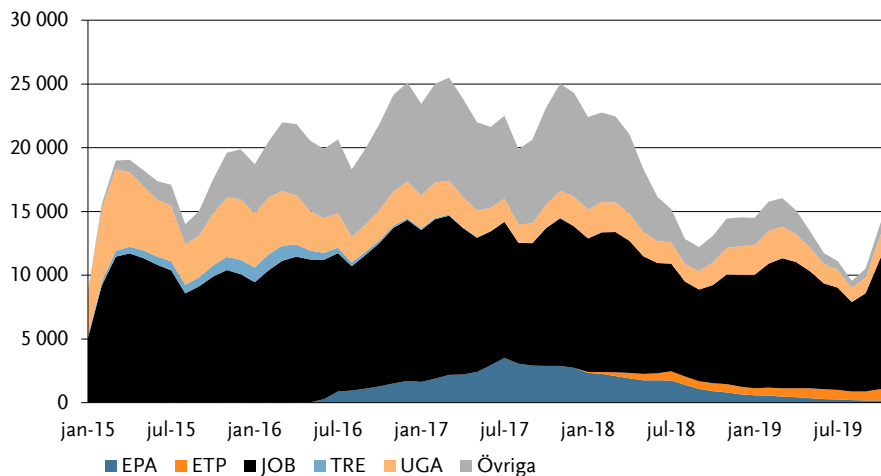
⁷ I övervägandet om huruvida en leverantör vill etablera sig i ett område antar vi att följande ingår: f_x [Potentiella deltagare (+), pris per deltagare (+), fasta kostnader för etablering (-), rörliga kostnader per deltagare (-)], där tecknet (+/-) visar om variabeln antas påverka positivt eller negativt. Ekvationen går givetvis att bygga ut med förväntningar om leverantörens egen kapacitet i förhållande till sammansättningen av potentiella deltagare. Men en sådan utbyggnad fyller inget syfte i frågan som analyseras.

3 Deltagarna

3.1 Deltagarsammansättningen över tid

Deltagarsammansättningen har förändrats inom STOM över tid (se figur A9). När STOM var ett relativt nytt program kom deltagarna mestadels från Jobb- och utvecklingsgarantin (JOB) och Jobbgaranti för ungdomar (UGA). Men över tid har sammansättningen förändrats. När STOM var som störst (under 2016 och 2017) kom en stor andel av deltagarna från UGA eller utanför garantierna (omkring 40 procent). Denna grupp är att betrakta som mer lättmatchade än andra eftersom de i genomsnitt har varit arbetslösa kortare tid och även har en förväntat kortare tid tills de finner arbete eller påbörjar studier. Men från våren 2018 minskade denna grupp betydligt. I slutet av 2019 utgjorde denna grupp (UGA och övriga utanför garantierna) endast omkring 20 procent av deltagarna. Lejonparten av programdeltagarna har alltid kommit från JOB men under det senaste året har detta blivit ännu mer framträdande.

Figur A9 Deltagarsammansättning



Anm.: EPA = Etableringsplan, ETP = Etableringsprogrammet, JOB = Jobb- och utvecklingsgarantin, TRE = Sysselsättningsfasen (fas 3), UGA = Jobbgarantin för ungdomar, Övriga = Övriga utanför garantierna eller etableringen.

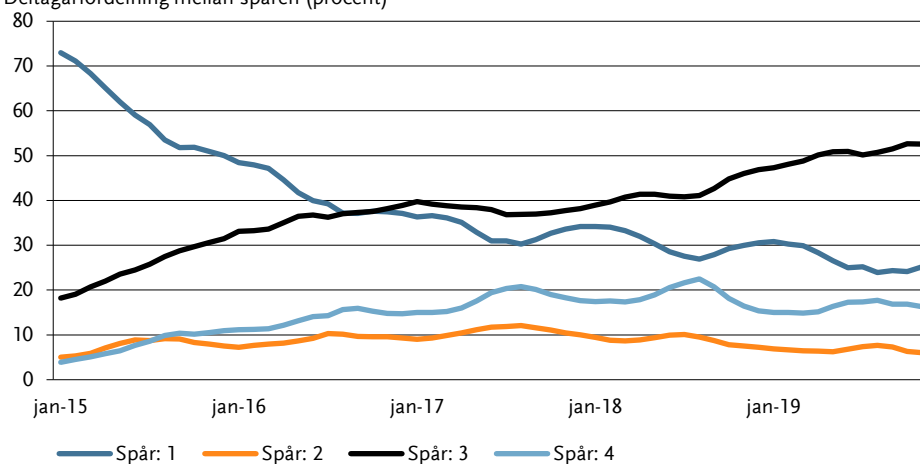
Källa: Arbetsförmedlingen.

Att deltagarsammansättningen har förändrats mot fler arbetssökande med större svårigheter att hitta jobb ska hanteras genom de fyra olika spårerna. Den snabba sammansättningsförändring som startade under våren 2018 borde följas av att andelen i spår 1 minskar. Det övre diagrammet i figur A10 bekräftar att andelen i spår 1 har minskat sedan våren 2018 men den utvecklingen är snarare en fortsättning på en redan tidigare minskande andel som tog fart redan i början av systemets lansering. Det är främst spår 3 som under denna period ökat på bekostnad av spår 1. Men även spår 4 har ökat (men främst under de två första åren).

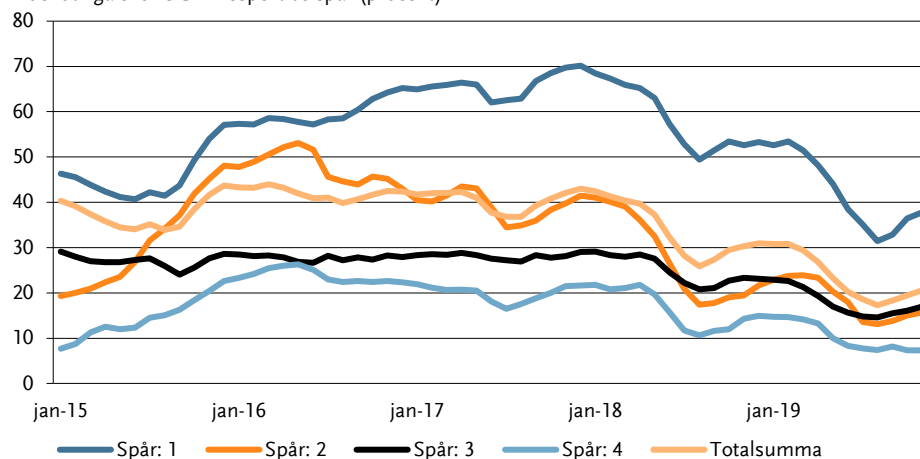
I det nedre diagrammet i figur A10 har vi beräknat andelen deltagare som kommer från UGA eller utanför garantierna och etableringen. Om den förändrade deltagarsammansättningen till fullo hade hanterats av storleken på de olika spåren borde sammansättningen inom spåren vara relativt oförändrade. Men även om de relativa storlekarna på spåren har förändrats har ändå deltagarfördelningen inom spår 1 blivit allt mindre fördelaktig, särskilt i början av 2018 då den sjönk från knappt 70 procent till knappt 40 procent under 2019. Den allt mer svårmatchade deltagarsammansättningen har alltså inte kompenserats tillräckligt genom spårssystemet.

Figur A10 Spårfördelning och deltagarsammansättning

Deltagarfördelning mellan spåren (procent)



Andel övriga eller UGA i respektive spår (procent)

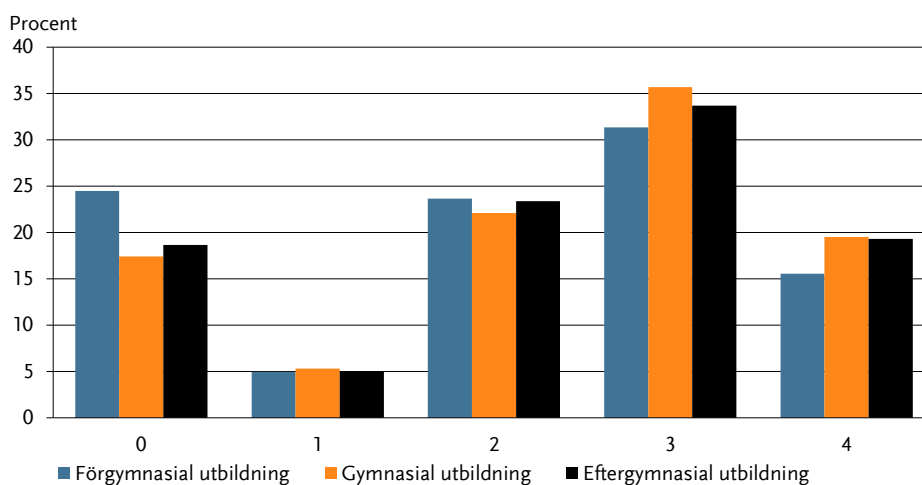


Källa: Arbetsförmedlingens webbplats.

3.2 Beskrivning av deltagarna uppdelat på rating

I detta avsnitt beskrivs om det finns skillnader i olika bakgrundsvariabler såsom utbildningsnivå för deltagare i STOM när det gäller valet av leverantör uppdelat på rating. Beskrivningen gäller främst perioden med fyra olika ratingnivåer som infördes i december 2018. Inledningsvis visas skillnader mellan olika utbildningsnivåer uppdelat på leverantörernas rating i figur A11. Med leverantörer avses här egentligen leverantörens avtal, eller leverantörer per leveransområden, eftersom det är på denna nivå ratingen beräknas.

Figur A11 Deltagares val av leverantör uppdelat på utbildningsnivå och leverantörens rating

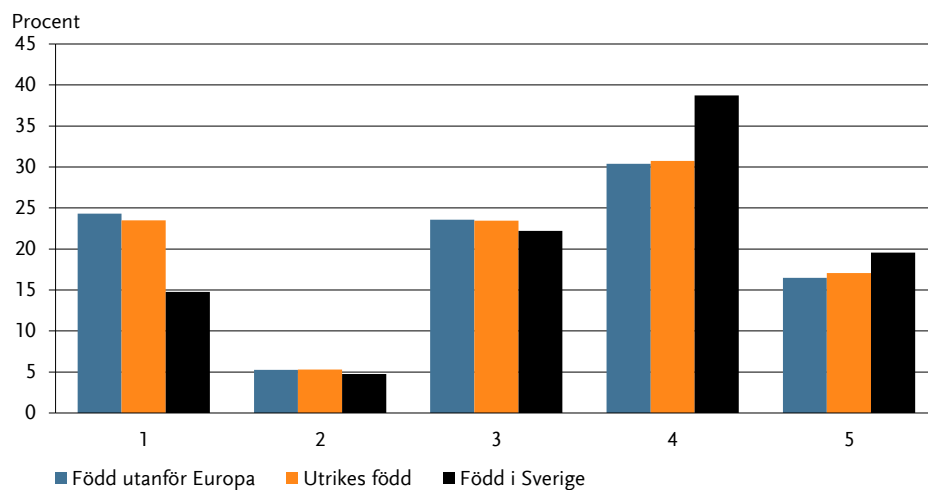


Källor: Arbetsförmedlingen och Statistiska centralbyrån.

Det finns en tendens att deltagare med endast förgymnasial utbildning i något högre utsträckning väljer leverantörer som inte har någon rating jämfört med deltagare som har högre utbildningsnivå. Samtidigt tycks deltagare med förgymnasial utbildning i något lägre utsträckning välja leverantörer med 3 och 4 som rating. I figur A12 undersöks om det finns några skillnader i deltagarnas val när det gäller var deltagarna är födda. Utrikes födda deltagare och delgruppen som är födda utanför Europa väljer i jämförelse med deltagare som är födda i Sverige oftare leverantörer som inte har någon rating. Samtidigt väljer dessa grupper mindre ofta de leverantörer som har tre respektive fyra i rating.

Analysen visar också att deltagare som har varit arbetslösa längre tid innan deltagande oftare väljer leverantörer utan någon rating och samtidigt mer sällan väljer leverantörer med den högsta ratingen. Vi ser ingen skillnad i valen av leverantör när det gäller leverantörernas rating när vi gör uppdelningen på kön eller ålder. Dock tycks deltagare boende i tätort oftare välja leverantörer med högre rating än de som inte bor i en tätort.

Figur A12 Deltagarnas val av leverantör uppdelat på födelseland och leverantörens rating



Anm.: Gruppen Född utanför Europa är en delmängd av Utrikes födda.

Källor: Arbetsförmedlingen och Statistiska centralbyrån.

4 Ratingen

Arbetsförmedlingens ratingmodell visar vilka leverantörer⁸ som lyckats bäst med att stötta deltagare till arbete eller studier.⁹ Ratingen rangordnar leverantörerna (avtalen) på STOM-marknaden från 1 till 4. Exempelvis ges 4 stjärnor till de leverantörer som har presterat bäst i landet givet deras förutsättningar. För att säkerställa beräkningarnas tillförlitlighet måste leverantörerna haft minst 18 deltagare i leveransområdet som *startade* insatsen stöd och matchning under året innan beräkningen utförs. Exempelvis krävs det att minst 18 deltagare påbörjade insatsen mellan september 2017 och augusti 2018 för att leverantören i det leveransområdet ska få en rating publicerad i september 2019. Med andra ord är det deltagare från STOM för 13 till 24 månader sedan som utgör underlaget för beräkningarna.

Grundtanken med ratingen är att leverantörernas prestation inte bara ska värdesättas utifrån hur många de får ut i arbete eller studier. Man behöver även väga in hur svårmatchade deltagarna hos en leverantör bedöms vara. Kombinationen av de två, utflöde och förutsättningar, avgör hur väl en leverantör har lyckats i jämförelse med andra leverantörer. Om exempelvis en leverantör oftast får relativt svårmatchade individer korregerar ratingmodellen för detta så att leverantörerna tävlar om att vara bäst på lika villkor. Det betyder i praktiken att leverantör A som får ut arbetssökande i arbete i snabbare takt än leverantör B ändå kan få lägre rating om leverantör B arbetar med grupper som enligt tidigare statistiskt mönster är mer svårmatchade.

Ratingmodellen kan beskrivas i fyra huvudsakliga steg:

1. Skatta individens risk för långtidsarbetslöshet med en logistisk regression.
2. Skatta sannolikheten att nå R2 för de 10 deciler av risk för långtidsarbetslöshet som skattades i steg 1. Denna sannolikhet benämns som profileringsvärdet.
3. Beräkna score-värden:
 - a. Beräkna residualen från steg 2: Profileringsvärdet, resultatet av steg 2, jämförs med det faktiska utfallet (R2).
 - b. Beräkna ett medelvärde av alla individuella residualer för varje leverantör.
4. Rangordna leverantörerna (ratingen) baserat på leverantörernas genomsnittliga score-värden.

I faktarutan nedan beskrivs modellen mer utförligt.

⁸ Med leverantörer avses här egentligen avtalen, eller leverantörer per leveransområden, eftersom det är på denna nivå ratingen beräknas.

⁹ Kapitlet är baserat på Arbetsförmedlingen, *Beskrivning av hur ratingen tas fram i tjänsten Arbetsförmedlingens kundval – Stöd och matchning*; Arbetsförmedlingen, *Beskrivning av framtagningen av ratingvärden för Stöd och matchning (STOM)*, 2017-08-29, Arbetsförmedlingen.

Ratingmodellen

1. Skatta risken för långtidsarbetslöshet

Den statistiska grundmodellen utgår från personer som har haft sin uppväxt och levt större delen av sitt liv i Sverige. Den följer det så kallade "bedömningsstödet" som har använts i Arbetsförmedlingens verksamhet sedan 2012. Sedan december 2016 bygger ratingvärdet på deltagare i alla spår (1 till 4); tidigare beräknades det bara för spår 1. Eftersom statistiska uppgifter rörande nyanlända eller personer i etableringsprogrammet (särskilt deras arbetsmarknadshistorik) är mycket knapphändiga eller saknas helt skattas en separat modell för denna grupp. Dessutom beräknas modellen separat för män och kvinnor.

I den allmänna profileringsmodellen som skattas på alla utom nyanlända och etableringsprogrammet ingår följande:

- födelseland uppdelad i inrikes född, Nordisk, Europa samt utanför Europa
- inträdesmånad i STOM
- indikator på funktionsnedsättning
- ålderskategorier
- indikator på medlemskap i a-kassa eller ej
- utbildningskategorier (förgymnasial, gymnasial, eftergymnasial)
- indikator för om personen har erfarenhet i sökt yrke
- uppdelning av sökt yrke i olika yrkesklasser utifrån SSYK-koder (Standard för svensk yrkesklassificering)
- indikator för senaste år en person varit inskriven som föregick nuvarande inskrivningsperiod
- uppdelning i tidigare arbetslöshetshistorik, dvs. antal dagar i arbetslöshet innan STOM-perioden
- arbetslösheten i kommunen där personen har varit bosatt.

Den beroende variabeln i regressionsmodellen är den binära variabeln långtidsarbetslös (definierad som minst 6 månader i en given inskrivningsperiod). De skattade koefficienterna från den logistiska modellen används sedan tillsammans med STOM-deltagarnas egenskaper (variablerna födelseland, ålder, sökt yrke etcetera.) för att beräkna individuell sannolikhet att hamna i långtidsarbetslöshet.

2. Skatta sannolikheten att nå R2

Eftersom att resultaten från steg 1 avser risken för långtidsarbetslöshet behövs ett ytterligare steg för att överföra resultaten till utfallet R2. Detta kan ses som det sista steget i den statistiska profileringen: Utifrån den individuella predikterade sannolikheten att hamna i långtidsarbetslöshet (som beräknades i steg 1) delas deltagarna in i deciler (10 lika stora rangordnade grupper). För varje decil och kön skattas en OLS-regression på utfallet R2 med kontroller för boendelän och spår som de deltar i.

Skattningarna genomförs endast för de individer som befinner sig inom uppföljningsperioden, det vill säga som har påbörjat och inte avbrutit tjänsten för något annat än arbete eller studier. Uppföljningsperioden påbörjas minst 13 och högst 15 månader efter programstart och pågår till 24 månader efter programstart:

9 (tjänstens maximala längd) + 0 till 2 (placeringsperiod till anställning) + 4 (minimal anställningstid) = 15 till 13 månader beroende på hur lång placeringsperioden till anställningen blev. Med placeringsperiod avses att anställningen måste ha tillträtts senast efter 2 månader räknad från tiden hos leverantören för att det lyckade utfallet ska hänföras till leverantören (så kallat R1). Sedan krävs minst 4 månaders anställning för att Arbetsförmedlingen ska ge ett godkänt R2.

3. Beräkna score-värden

När varje deltagare väl har en beräknad sannolikhet jämförs detta mot det faktiska utfallet (faktiskt utfall – profileringsvärde), det vill säga residualen från steg 2 beräknas för varje individ. Sedan beräknas ett medelvärde av individernas residualer för varje leverantör, ett så kallat score-värde. Score-värdet uttrycker då hur stor andel som gick till R2 för en leverantör jämfört med vad man kunde förvänta sig utifrån modellen.

4. Rangordna leverantörerna (ratingen) baserat på leverantörernas score-värde

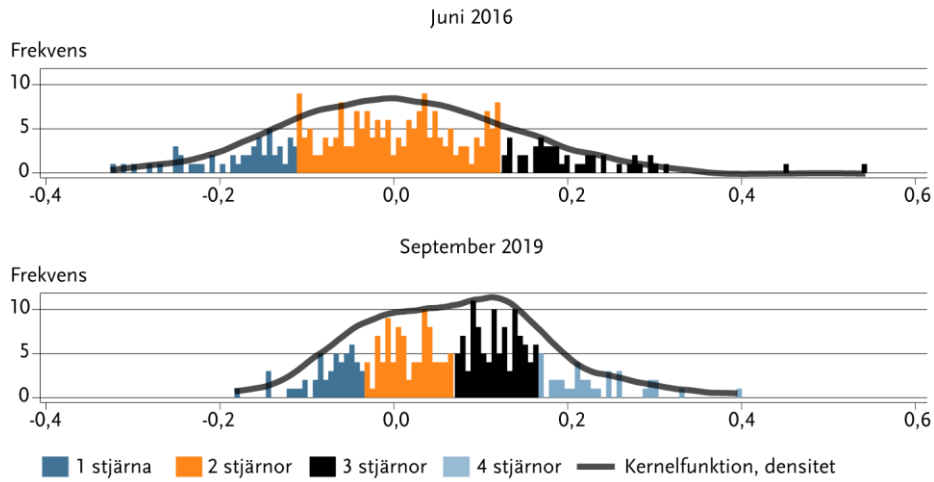
Alla leverantörers scorevärden rangordnas och delas in i 4 grupper (tidigare 3 grupper):

- 1: score < -1 standardavvikelse
- 2: -1 standardavvikelse =< score < medelvärdet
- 3: medelvärde =< score < 1 standardavvikelse
- 4: score => 1 standardavvikelse

I figur A13 illustreras fördelningen av score-värden i den första estimeringsrundan (juni 2016) och i estimeringsrunda 14 (september 2019). När ratingmodellen togs fram blev medelvärdet av leverantörernas score-värden 0,01 och standardavvikelsen 0,13 vilket motsvarade cirka 13 procentenheters skillnad i sannolikhet att få R2 under uppföljningstiden. En genomsnittlig arbetssökande som valde en leverantör den estimeringsrundan med 3 stjärnor kunde alltså förvänta sig en standardavvikelse (det vill säga 13 procentenheter) högre sannolikhet att få jobb eller utbildning jämfört med en genomsnittlig leverantör. Exakt vilka leverantörer som är helt genomsnittliga framgår dock inte av ratingsystemet, bara att de finns i gruppen med två stjärnor. Det är fullt möjligt att en individ väljer mellan två leverantörer som ligger precis intill gränsen för 2 respektive 3 stjärnor och att valet mellan dem således inte är särskilt avgörande för individens jobbchans. Under årens lopp har medelvärdet för score-värdet ökat och standardavvikelsen minskat något. I figur A13 framgår att det främst är de lägst presterande som inte längre finns kvar. För den senaste estimeringsrundan kunde en genomsnittlig arbetssökande som valde en topprankad leverantör förvänta sig i genomsnitt 16 procentenheter högre sannolikhet att finna jobb eller utbildning under uppföljningsperioden jämfört med om hen valde en genomsnittlig leverantör (tabell A2).¹⁰ Men även om skillnaden nu minskat mellan leverantörer med olika stjärnor var det förväntade resultatet jämfört med modellen ändå oförändrat på cirka 15 procentenheter (medelvärdet + standardavvikelsen). Att medelvärdet har ökat kan bero på att populationen som profileringsmodellen utgår från och den slutliga populationen som score-värdet beräknas på kan skilja sig åt på grund av manuella efterhandsjusteringar av Arbetsförmedlingen till följd av kvalitetsbrister i indata.¹¹ En annan möjlig, mer teknisk, anledning har att göra med att prediktionerna görs på individnivå men att beräkningen som leder till score-värde på leverantörsnivå inte beaktar hur många individer prediktionen beräknades på. Lite enkelt uttryckt så blir inte medelvärdet av flera medelvärden samma sak som ett medelvärde av alla ingående tal. Hade medelvärdet viktats med antal individer i beräkningen hade medelvärdet hamnat närmare noll, allt annat lika.

¹⁰ För estimeringsrunda 14 hade en topp-rankad leverantör i genomsnitt 16 procentenheter (0,23-0,07=0,16) bättre resultat än medelvärdet för samma estimeringsrunda.

¹¹ E-post från företrädare för Arbetsförmedlingen, 2020-02-07.

Figur A13 Fördelning av score-värden

Källa: Arbetsförmedlingen.

Tabell A2 Score-värden fördelat på estimeringsrundor.

Score-värde: Skillnad i andel till R2 mellan en prediktiv modell och faktiskt utfall

Estimeringsrunda	Medelvärde	Standardavvikelse	Antal	Medelvärde för rating			
				1	2	3	4
1 (2016-06)	0,01	0,13	279	-0,18	0,00	0,21	
2* (2016-09)							
3 (2016-12)	0,00	0,12	342	-0,15	0,01	0,18	
4 (2017-03)	0,04	0,12	366	-0,12	0,04	0,21	
5 (2017-06)	0,03	0,13	376	-0,14	0,04	0,21	
6 (2017-09)	0,03	0,12	407	-0,14	0,03	0,20	
7 (2017-12)	0,05	0,11	386	-0,11	0,05	0,21	
8 (2018-03)	0,04	0,12	395	-0,13	0,04	0,20	
9 (2018-06)	0,05	0,09	343	-0,08	0,04	0,18	
10 (2018-09)	0,06	0,10	310	-0,07	0,05	0,22	
11 (2018-12)	0,07	0,10	284	-0,10	0,03	0,12	0,26
12 (2019-03)	0,07	0,09	273	-0,06	0,03	0,11	0,22
13 (2019-06)	0,07	0,09	264	-0,08	0,02	0,11	0,24
14 (2019-09)	0,07	0,10	275	-0,07	0,02	0,12	0,23
Totalt	0,04	0,11	4 300				

Anm.: * Uppgifter för estimeringsrunda 2 saknas. Score-värde anger hur väl en leverantör lyckades med att nå R2 i förhållande till en statistisk modell. Exempelvis betyder ett medelvärde på 0,07 att leverantörerna i genomsnitt fick 7 procentenheter fler i arbete än vad den statistiska modellen predikerade. En standardavvikelse på 0,10 betyder således för en individ att de har minst 10 procentenheters högre chans att nå godkänt R2 om de väljer en leverantör med högst rating.

Källa: Arbetsförmedlingen.

5 Incitament för effektivt beteende hos leverantörer

Lönar det sig att vara högpresterande och straffar det sig att underpresteras? Detta kapitel undersöker systemöverlevnad bland leverantörernas avtal totalt och fördelat på rating, samt hur ratingen påverkar deltagarinflödet för leverantörerna. Även sambandet mellan leverantörsstorlek och rating undersöks. Kapitlet avslutas med en granskning av kvaliteten i R2. Detta görs dels genom att följa upp kortsiktigt mot uppgifter hos Skatteverket, dels genom en jämförelse av arbetsmarknadsutfallet för individen cirka ett år efter godkänd R2 mellan leverantörer med olika rating.

Studier av ratingsystem har gjorts tidigare, till exempel ratingsystemet till Arbetsförmedlingens ”jobbcoach”, som påminner om ratingsystemet i dagens STOM. Studien fann att bästa rating fick fler deltagare men att grupper med i genomsnitt relativt stora svårigheter att hitta jobb såsom personer födda utanför Europa och lågutbildade inte reagerade på systemet i lika hög utsträckning som övriga.¹² Även Arbetsförmedlingen har gjort en mindre opublicerad studie av STOM:s rating till Arbetsmarknadsutredningen.¹³ Huvudresultatet från den studien var att en ökad rating leder till cirka 20 procent fler deltagare.

5.1 Datamaterialet och ratingen

Ratingen görs en gång var tredje månad. Exempelvis beräknas ratingen i mars och gäller fram till maj, sedan görs en ny beräkning för juni och så vidare. I december 2018 utökades antal stjärnor från 3 till 4. Detta gjordes genom att helt enkelt dela gruppen med 2 stjärnor mitt i tu. Således är datamaterialet indelat i två tydliga perioder, före och efter ratingutökningen. Den första perioden omfattar således drygt 30 000 observationer (varav nästan 20 000 med uppgifter om rating) fördelat på cirka 1 200 leverantörer (drygt 1 000 med uppgifter om rating) över 48 månader (31 månader med uppgifter om rating) medan den andra perioden omfattar knappt 5 000 observationer fördelat på drygt 580 leverantörer över 10 månader. Notera än en gång att en leverantör räknas som en unik leverantör per område. Det betyder att en organisation som deltar på STOM-marknaden kan räknas som olika leverantörer om organisationen har kontor i olika leveransområden. För varje organisationsnummer kan det alltså finnas flera avtal (vilket i Riksrevisionens data betecknas som KA-nummer). Det är främst på denna nivå som fokus och beräkningarna ligger på, det vill säga leverantör per leveransområde.

¹² Gerdes, ”Does Performance Information Affect Job Seekers in Selecting Private Providers in Voucher-Based ALMP Programs? | IZA - Institute of Labor Economics”, 2015.

¹³ SOU 2019:3, *Effektivt, tydligt och träffsäkert – det statliga åtagandet för framtidens arbetsmarknad*, 2019.

Tabell A3 beskriver antalet förstagångsdeltagare¹⁴ över de två perioderna och fördelat på antal ratingstjärnor. Tabellen avser endast perioden med tillgång till ratingvärden och för leverantörer med minst en förstagångsdeltagare. Det finns ett tydligt positivt samband mellan antalet stjärnor och antalet förstagångsdeltagare oavsett vilken period som studeras. Exempelvis fick leverantörer i den första perioden med 3 stjärnor drygt 2 fler förstagångsdeltagare per månad jämfört med leverantörer med 2 stjärnor (motsvarar 24 procent fler). I den andra perioden hade leverantörer med 4 stjärnor knappt 4 fler förstagångsdeltagare per månad jämfört med leverantörer med 3 stjärnor (motsvarar 51 procent fler). Att jämföra mellan perioderna blir dock lite vanskligt då de har delats in i olika många ratingnivåer.

Tabell A3 Genomsnittligt inflöde av nya förstagångsdeltagare till leverantörer per månad, fördelat på rating. Leverantörer med minst en förstagångsdeltagare. Genomsnitt per månad mellan juni 2016 och september 2019

	Förstagångsdeltagare			Observationer
	Medelvärde	Standardavvikelse	Max	Antal
Juni 2016–november 2018				
Inga stjärnor	4,6	5	78	6 167
1	6,5	8	103	1 667
2	10,1	11,4	121	5 995
3	12,5	13,4	111	1 869
Totalt	7,9	9,8	121	15 698
December 2018–september 2019				
Inga stjärnor	3,2	3	22	1 315
1	4	4,5	46	262
2	5,9	5,8	48	817
3	7,7	7,1	52	909
4	11,6	12,2	73	323
Totalt	5,7	6,7	73	3 626

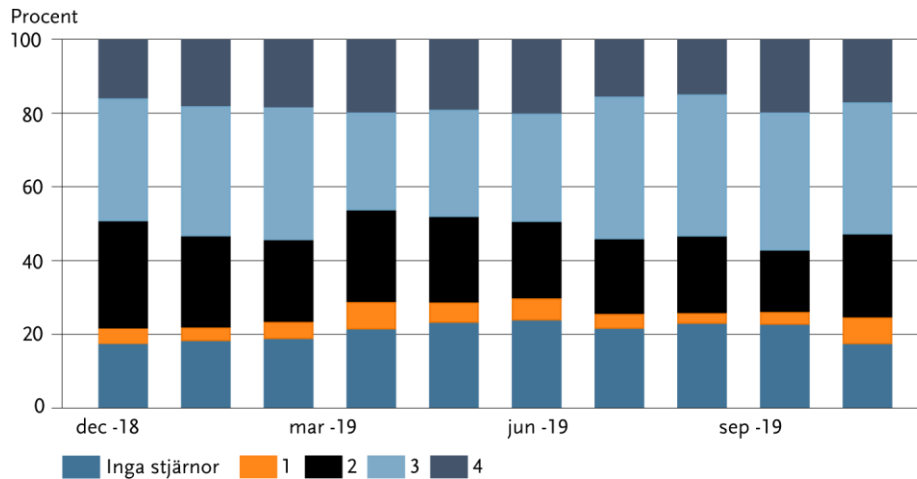
Källa: Arbetsförmedlingen.

Att benägenheten bland förstagångsdeltagare att gå till de lägst rankade leverantörerna är lägre gör även att andelen av alla förstagångsdeltagare som går till lågt rankade leverantörer blir låg. För samma period (december 2018–september 2019) var det i genomsnitt 5 procent av alla förstagångsdeltagarna som valde en leverantör med

¹⁴ I utredningen, SOU 2019:3, framgick att de som inte var förstagångsväljare ofta inte valde leverantör aktivt (endast cirka 20 procent gjorde ett aktivt val). Därför gjordes större delen av analysen på förstagångsdeltagare. Riksrevisionen valde därför att analysera samma grupp. Det har dock senare visat sig i kommunikationen med Arbetsförmedlingen att "omväljare" har samma grad av aktivitet i valet när personer som får förlängda beslut inte längre räknas som "omväljare". Riksrevisionen har testat att inkludera alla inflödande deltagare till leverantörerna (där förlängningar inte räknas som en ny period). Inkluderingen gör mycket lite för parameterestimaten. Om något, blir de marginellt starkare.

1 stjärna (se figur A14). Även om andelen leverantörer med 4 stjärnor är lika stor som andelen med 1 stjärna är således andelen förstagångsdeltagare som går till de högst rankade leverantörerna betydligt högre.

Figur A14 Andelen förstagångsdeltagare fördelat på leverantörsrating och tid



Källa: Arbetsförmedlingen.

Under de första månaderna efter att en leverantör etablerat sig i ett leveransområde kommer de att sakna rating i minst 13 månader (på grund av hur regelverket ser ut), därför blir även den genomsnittliga övergångssannolikheten hög. Men från och med att de fått sin första rating sjunker sannolikheten för att skifta rating till mellan 52 och 63 procent beroende på vilken rating de hade (se tabell A4). Exempelvis var sannolikheten att behålla sin topprating mellan två perioder 63 procent medan sannolikheten att stanna i den lägsta ratingen var 52 procent. Tabell A4 visar även att det är relativt ovanligt att förlora (eller vinna) fler än en stjärna mellan två publiceringstillfällen. För leverantörer med 3 stjärnor förlorade (vann) i genomsnitt 22 (10) procent en stjärna medan endast 2 procent förlorade 2 stjärnor.

Tabell A4 Övergångssannolikheter mellan två publiceringstillfällen, dec 2018–sep 2019

t/t+1	Ingen rating	Rating 1	Rating 2	Rating 3	Rating 4
Ingen rating	90%	3%	3%	2%	2%
Rating 1	21%	52%	21%	6%	0%
Rating 2	3%	16%	59%	21%	1%
Rating 3	2%	2%	22%	63%	10%
Rating 4	1%	1%	5%	31%	63%

Källa: Arbetsförmedlingen.

Ett annat mått på ratingens stabilitet är den genomsnittliga andelen tid en leverantör har en viss rating. Kolumnen ”Inom leverantören” i tabell A5 visar exempelvis att om en leverantör någon gång når 4 stjärnor så har de tillbringat 49 procent av sin tid (i den undersökta perioden) i det tillståndet. Givet att en leverantör alltid haft en rating under perioden visar detta stabilitetsmått att rating 2 och 3 är de mest stabila nivåerna och att rating 1 är den mest instabila nivån.

Tabell A5 Ratingens variation mellan och inom leverantörer, dec 2018–sep 2019

Rating	Totalt		Inom leverantören
	Antal	Procent	Procent
Ingen rating	926	46	88
1	152	8	46
2	393	19	55
3	408	20	57
4	143	7	49
Totalt	2 022	100	66

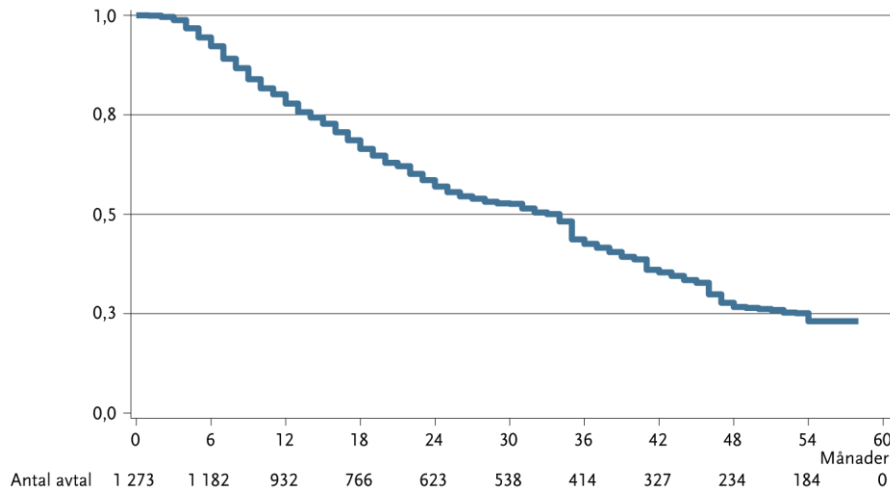
Källa: Arbetsförmedlingen.

5.2 Lönar det sig med en hög rating?

I detta kapitel undersöker vi hur leverantörer påverkas av den rating de har. Först undersöker vi hur ratingen påverkar systemöverlevnaden och sedan hur ratingen påverkar flödet av nya deltagare till leverantörerna.

5.2.1 Systemöverlevnad

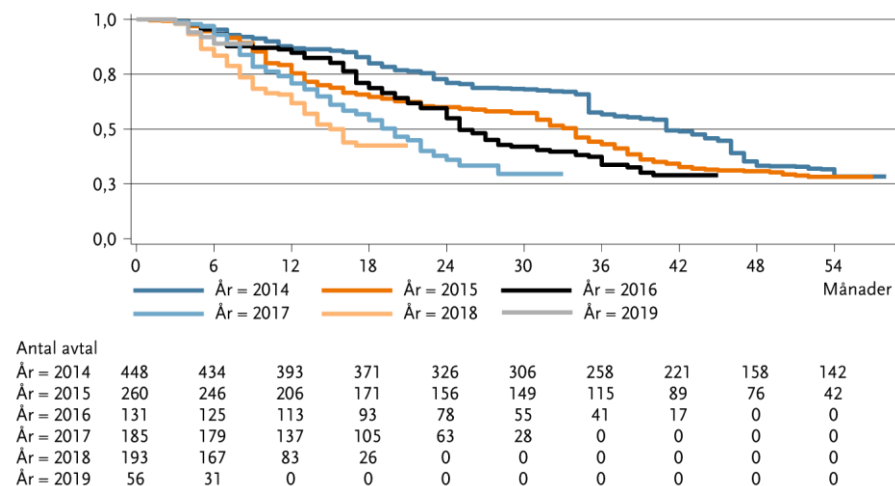
Figur A15 är ett så kallat Kaplan–Meier-diagram som visar systemöverlevnaden för leverantörernas avtal. Det framkommer här att hälften överlever minst 33 månader och att endast 25 procent har klarat sig i systemet minst 53 månader. Utslagningstakten är ganska jämn över tid.

Figur A15 Leverantörsöverlevnad i STOM-systemet

Anm.: Linjen visar andelen aktiva avtal från och med att de påbörjades. Riksrevisionen har uppgifter om systemet från december 2014 till september 2019.

Källor: Arbetsförmedlingen och egna beräkningar.

Olika inträdesår på STOM-marknaden har gett olika överlevnadskurvor i systemet. Det är främst de som tidigt klev in på marknaden som har klarat sig längst i systemet. För de avtal som startade under 2014 var cirka 85 procent kvar efter 18 månader medan endast 50 procent av de som startade 2017 var kvar efter lika lång tid. En viktig delförklaring till detta är att antalet deltagare till STOM minskade kraftigt från och med våren 2018 ända till mitten av 2019. För de som etablerade sig i ett leveransområde under 2019 verkar utvecklingen hittills följt den genomsnittliga utvecklingen.

Figur A16 Systemöverlevnad fördelat på inträdesår

Källor: Arbetsförmedlingen och egna beräkningar.

Hur påverkar då ratingen risken för att slås ut? Eftersom data både är vänstertrunkerad och högercensurerad samtidigt som ratingen är tidsvarierande lämpar sig en diskret tidsmodell bäst för en sådan analys.^{15,16} Tidsberoendet kan modelleras antingen stegvis¹⁷ eller genom någon specificerad funktionsform av tidens påverkan på risken att slås ut ur systemet. I den diskreta tidsmodellen görs antagandet om att risken¹⁸ är proportionell till tidsberoendet. Det betyder i praktiken att koefficienterna till de oberoende variablerna skiftar hela riskfunktionens profil upp eller ner. Det innebär alltså att vi antar att effekten av att ha exempelvis 2 stjärnor under de första fem månaderna har samma (proportionella) effekt på risken som att ha 2 stjärnor efter exempelvis 30 månader. Men det är möjligt att leverantörerna inte påverkas särskilt mycket av ratingen de första månaderna i systemet. Det är exempelvis troligt att leverantörerna försöker att klara verksamheten i en uppbyggnadsfas oavsett rating innan den låga ratingen börjar ta ut sin rätt. I vår analys är detta dock ett mindre problem eftersom vi främst är intresserade av den genomsnittliga effekten av ratingen snarare än av hur den slår igenom i olika tidsskeden.

Tabell A6 sammanfattar 6 diskreta tidsmodeller där 1 stjärna (harmoniserad rating¹⁹) utgör referenspunkten för övrig rating. Modell 1 visar exempelvis att risken för att leverantörer med 2 stjärnor slås ut är cirka 18 procent av risken för att en leverantör med 1 stjärna ska slås ut. Motsvarande siffra för leverantörer med 3 stjärnor är cirka 11 procent av risken oavsett tid i systemet. Det finns dock ingen skillnad i risk mellan att sakna rating och att ha 1 stjärna. I modell 2 inkluderas även antalet nya deltagare för varje månad. Den exponerade koefficienten på 0,72 visar att fler deltagare är associerat med en 28 procents lägre risk att slås ut. Samtidigt minskar den uppmätta riskreduceringen av ratingen, från cirka 82 till 73 procent för 2 stjärnor och från 89 till 75 procent för 3 stjärnor. Men eftersom inflödet av nya deltagare delvis är en funktion av hög rating är det inte helt entydigt att modell 2 är att föredra, men den redovisas ändå här för transparensens skull. Resten av modellerna testar hur resultaten förändras när vi kontrollerar för andra variabler istället. Slutsatsen är att förhållandet mellan ratingen och risken att lämna STOM är tämligen robust, det vill säga fler stjärnor ökar sannolikheten att leverantören klarar sig längre tid i systemet.

¹⁵ Vänstertrunkering beror på att uppgifterna om leverantörernas rating endast finns tillgängliga från juni 2016 och framåt trots att många leverantörer trädde in på STOM-marknaden innan dess. Högercensureringen är en naturlig konsekvens av att vi observerar ett system som fortfarande är igång. Det innebär att vi inte vet vilka leverantörer som deltog för sista gången i september (eftersom vi inte har några uppgifter om oktober). För att kunna inkludera tidsvarierande variabler såsom rating kan detta estimeras i en diskret tidsmodell med en *komplementär log-log* länk. Andra kända länkfunktioner för att estimeras modeller med binära beroende variabler är exempelvis de mer kända logit-länken och probit-länken. Den *komplementära log-log*-länken lämpar sig väl för när den beroende binära variabeln sällan antar värdet 1. Modellen handskas även med vänstertrunkeringen och högercensureringen och är rätt lik Cox-regressioner. Det bärande antagandet i modellen är att variablerna (exempelvis ratingen) påverkar grundrisken (*baseline hazard*) proportionerligt över hela analys tiden.

¹⁶ För mer om diskret-tidsmodellering, se Jenkins, "Easy Estimation Methods for Discrete-Time Duration Models", 1995; Allison, "Event History and Survival Analysis: Introduction", 2014.

¹⁷ Genom en s.k. *Piecewise constant hazard*. Det vill säga olika risker för varje månad men konstant risk (*hazard*) inom respektive månad. Men om inget avtal avslutades en viss månad antar vi att risken är densamma som föregående månad.

¹⁸ Även ofta kallat "*hazard*" men i denna rapport kommer vi att skriva *risken*.

¹⁹ Ratingen med fyra nivåer har kodats om till tre nivåer för att tillse att jämförelser över tid är möjliga.

Tabell A6 Diskreta tidsmodeller, relativ risk att lämna systemet

	(1)	(2)	(3)	(4)	(5)	(6)
Ingen rating	1,008	0,638**	1,000	1,125	1,433**	0,989
	(0,121)	(0,077)	(0,121)	(0,146)	(0,200)	(0,099)
1	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
2	0,183**	0,273**	0,185**	0,173**	0,199**	0,195**
	(0,023)	(0,035)	(0,023)	(0,022)	(0,029)	(0,023)
3	0,114**	0,245**	0,116**	0,099**	0,144**	0,123**
	(0,029)	(0,061)	(0,029)	(0,025)	(0,039)	(0,030)
Antal nya deltagare		0,719**				
		(0,024)				
Leverantörens storlek (antal avtal per leverantör)			0,994			
			(0,004)			
ln(t)						1,812**
						(0,118)
Konstant						0,008**
						(0,002)
Piecewise constant hazard	Ja	Ja	Ja	Ja	Ja	Nej
Kontroller för leveransområden	Nej	Nej	Nej	Ja	Nej	Nej
Kontroller för leverantörer	Nej	Nej	Nej	Nej	Ja	Nej
Observationer	23 593	23 593	23 593	23 548	20 975	23 593
BIC	5 919	5 512	5 927	6 418	6 668	5 655
Chi ²	8 010	6 174	8 048	.	.	295

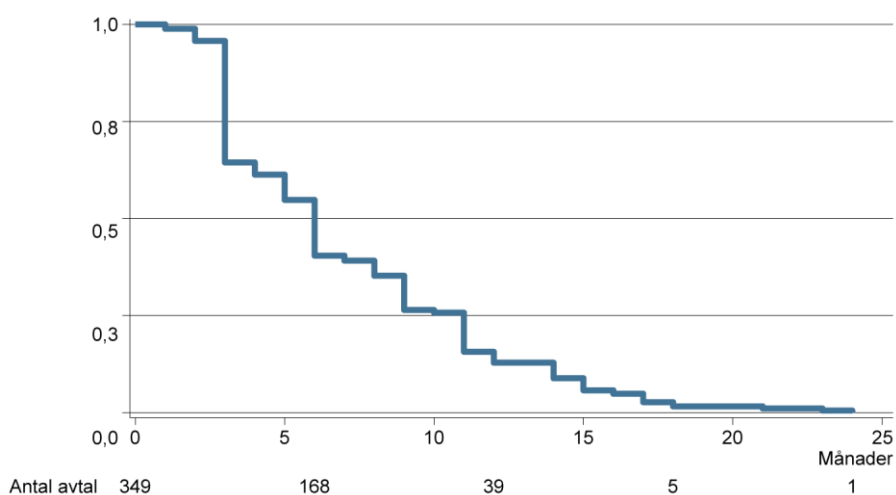
Anm.: I modell 6 är riskfunktionen specificerad till $\ln(t)$ istället för att tillåtas variera fritt mellan månaderna (s.k. parametrisk proportionell riskmodell). Koefficienterna är exponentierade (e^{β}) och återger således den relativa risken. Standardfelen är klustrade på avtal, * $p < 0,05$; ** $p < 0,01$.

Källor: Arbetsförmedlingen och egna beräkningar.

Systemet behåller inte lågt rankade avtal särskilt länge. Ett sätt att illustrera detta är att beräkna överlevnaden för de lägst rankade (1 stjärna) oavsett när deras period som lågrankad påbörjades, alltså oavsett om låg rating inträffade i och med deras systeminträde eller om de tappade stjärnor efter exempelvis 2 år. Både systemutträde och en förbättrad rating räknas här som slut på perioden. Om de däremot förlorar sin rating helt räknas det som att deras period som lågrankad fortsätter. Det är nämligen intressant att undersöka hur länge de lägst presterande avtalen kan verka i systemet utan att förbättras eller slås ut. Figur A17 visar att nästan hälften av avtalen som påbörjar en period med 1 stjärna avslutas genom att slås ut eller öka i rating efter 5 månader. Efter 12 månader har 87 procent slagits ut eller förbättrats. Den längsta sammanhängande perioden med 1 stjärna eller ingen rating var 24 månader (varav

18 av dessa månader helt saknade rating). Den längsta sammanhängande perioden med ett avtal med konstant 1 stjärna var 18 månader.²⁰ Notera att figur A17 inte är direkt jämförbar med tabell A6 eftersom diagrammet inte beaktar hur länge avtalen varit aktiva innan det fick sin rating, att utträde i detta fall inte bara är utträde ur systemet utan även förändrad rating samt att perioder utan rating skrivs över med senast kända rating.

Figur A17 Andel avtal vars senaste rating var 1 stjärna (saknad rating ersätts med 1 stjärna om avtalet hade 1 stjärna tidigare)



Anm.: Om leverantören förlorar sin rating används den senast fastställda ratingen. Om en leverantör exempelvis först får 1 stjärna och sedan förlorar sin rating räknas perioden utan rating som om den fortfarande hade 1 stjärna fram tills att en ny rating fastställs eller att leverantören lämnar systemet.

Källor: Arbetsförmedlingen och egna beräkningar.

5.2.2 Inflöde av nya deltagare

Metod

För att skatta effekten av ratingen på antalet nya deltagare vill vi kontrollera för sådant som inte är ett utfall av rating men som i sin tur kan påverka både rating och utfallet vi är intresserade av. Exempel på sådant kan vara geografiska faktorer och leverantörens profilering (i den mån profilen är oförändrad över tid). Därför skattas modeller med så kallade fixa effekter för avtal (leverantör–leveransområde) och tid (år–månad).

Utfallsvariabeln är (det logaritmerade) inflödet av förstagångsväljare. Anledningen till att vi begränsar inflödet till förstagångsväljare är att de till cirka 90 procent väljer leverantör aktivt.²¹ I vissa fall harmoniserar vi även ratingen över tid, det vill säga kodar om 4 nivåer till 3 nivåer, för att kunna utnyttja en längre period vilket exempelvis ger säkrare skattningar.

²⁰ Det fanns alltså tre avtal som hade en mycket lång period utan rating som både föregicks och efterföljdes av en kortare period med 1 stjärna.

²¹ När förlängningar inte räknas som omväljare är även andel av totalt inflöde av STOM-deltagare cirka 90 procent. Detta upptäcktes sent i analysprocessen men förändrar heller inte resultaten.

Resultat

Modell 1 i tabell A7 visar att det enkla linjära sambandet mellan harmoniserad rating²² och inflöde är 0,35, det vill säga cirka 40 procent mer inflöde för leverantörer med en ytterligare stjärna. I de linjära modellerna inkluderas inte avtal som saknar rating eftersom vi inte vill göra något antagande om ett linjärt förhållande mellan en stjärna och ingen alls.²³ Detta samband ändras knappt när tidsfixa effekter introduceras (modell 2). I modell 3 och 4 introduceras *random effects* och *fixed effects*. Båda estimatorerna sänker det skattade sambandet till 0,22 respektive 0,21.²⁴ Modell 4 ska tolkas som att när en leverantör ökar sin rating med en stjärna ökar inflödet av förstagångsväljare med cirka 23 procent ($100 * (\exp(0,21) - 1) \approx 23\%$). I modell 5 är förhållandet modellerat med en flexibel form där även avtal utan rating ingår (1 stjärna är referenskategori). Det är främst förflyttningar till den högsta ratingen som har gett ökat inflöde i genomsnitt över hela undersökningsperioden. Från 1 till 2 stjärnor ger cirka 10 procent fler deltagare medan 2 till 3 ger cirka 40 procent mer ($0,44 - 0,09 = 0,35 \approx 40$ procent). Vi finner inga statistiskt säkerställda skillnader mellan avsaknad av rating och en stjärna.

Tabell A7 Regressionstabell – Ratingens påverkan på (logaritmerat) inflöde av förstagångsväljare

	(1)	(2)	(3)	(4)	(5)
Rating harmoniserad	0,35**	0,36**	0,22**	0,21**	
	(0,04)	(0,04)	(0,03)	(0,03)	
Inga stjärnor					-0,03
					(0,04)
1					Ref.
					(.)
2					0,09**
					(0,03)
3					0,44**
					(0,05)
Konstant	1,05**	1,27**	1,41**	1,69**	1,69**
	(0,08)	(0,10)	(0,07)	(0,08)	(0,05)
Fixa tidseffekter	Nej	Ja	Ja	Ja	Ja
R ²	0,041	0,106		0,222	0,154

²² I dag finns det 4 stjärnor i ratingssystemet, innan december 2018 fanns det 3 stjärnor. När systemet gick från 3 till 4 stjärnor delades gruppen med 2 stjärnor itu. Därför går det att helt enkelt koda tillbaka det nya systemet så att nya systemets 2 och 3 blir 3, samt att 4 blir 3.

²³ Eftersom ratingen egentligen är en kategorivariabel kan det ses som problematiskt att anta att ratingen har en linjär funktionsform. Men även om antagandet om ett linjärt samband inte korrekt beskriver relationen mellan rating och utfall så förenklar det framställningen av resultaten när modellen utvecklas med olika typer av interaktioner. Tolkningen av det enkla parameterestimatet är också enkelt. Den fångar den genomsnittliga viktade förändringen av antalet förstagångsdeltagare vid en förändrad rating.

²⁴ Ett Hausman-test visar att modellerna 3 och 4 inte skiljer sig åt vilket innebär att båda estimatorerna är konsistenta men att *random effects* är mer effektiv (det vill säga bättre precision i skattningarna). Men det är endast i denna specifikation som testet favoriserar *random effects*. När vi delar upp urvalet på före och efter december 2018 eller specificerar en mer flexibel funktionsform favoriserar testet *fixa effekter*.

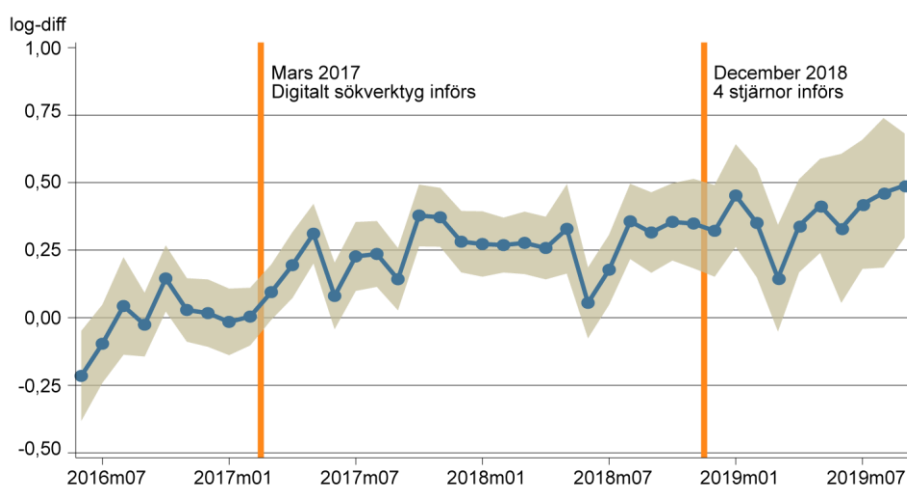
	(1)	(2)	(3)	(4)	(5)
Modell	ols	ols	re	fe	fe
Observationer	11 842	11 842	11 842	11 842	19 324

Anm.: Standardfel inom parenteser är klustrade på avtal. * $p < 0,05$; ** $p < 0,01$.

Källor: Arbetsförmedlingen och egna beräkningar.

Figur A18 visar modell 4 från tabell A7 fast med interaktioner mellan rating och tidsdummys. Effekten av rating var obefintlig vid introduktionen av systemet. Men från och med mars 2017 då ett digitalt sökverktyg infördes uppkommer en gradvis ökande effekt över tid. Mellan juni 2017 och november 2017 var effekten till synes stabil över tid. Från och med december 2018 infördes 4 stjärnor i ratingen. Alla resonemang om förändring av effekten över tid kopplat till vissa tidpunkter är behäftade med osäkerhet. Vi vet nämligen inte hur effekten hade utvecklats sig i avsaknad av de olika förändringarna eftersom vi inte har någon kontrollgrupp att jämföra med. Därför kan vi aldrig med säkerhet säga att förändringen efter en viss tidpunkt berodde på regelförändringen vid tidpunkten även om det verkar troligt när utvecklingen innan förändringen var relativt jämn.

Figur A18 Linjär effekt av harmoniserad rating över tid (modell 4 med tidsinteraktioner)



Anm.: Modellen är skattad med fixa effekter för avtal och tid. Det gråskuggade området utgör ett 95-procentigt konfidensintervall beräknat med klustrade standardfel på avtal.

Källor: Arbetsförmedlingen och egna beräkningar.

I tabell A8 fortsätter vi att utveckla den föredragna modellen (med fixa effekter för avtal och tid) för att analysera hur övergången till 4 stjärnor påverkade inflödet av deltagare. Dessutom utelämnar vi systemets första år (juni 2016 till maj 2017) för att kunna göra mer rättvisande jämförelser mellan perioden med 3 stjärnor och 4 stjärnor. Modell 2 visar att effekten i före-perioden av en förändrad rating var cirka 13 procent (-0,14 log-diff) lägre än i efter-perioden. Dummys som indikerar före-perioden visar

i sin tur att antalet nya deltagare per månad och leverantör var högre i före-perioden.²⁵ Modell 3 är en linjär modell med fixa effekter skattad under perioden med fyra nivåer i ratingsystemet. Koefficienten på 0,20 innebär att när en leverantör får en ytterligare stjärna ökar deltagarinflödet, i genomsnitt, med cirka 22 procent. Modell 4 till 5 är flexibla modeller av ratingens påverkan på inflödet av deltagare uppdelat på före- och efterperioden. Här framgår att ratingeffekten främst uppstår vid de högre nivåerna. I efter-perioden ökar inflödet av deltagare med cirka 28 procent från 3 till 4 stjärnor.²⁶ Omräknat till antal deltagare blir det knappt 2 deltagare per månad. Effekten på antalet förstagångsdeltagare av en topprating jämfört med 1 stjärna var 55 procent respektive 68 procent i före-perioden respektive efter-perioden. Skillnaden i effekt av en topprating mellan de olika perioderna är signifikant (har testats i en och samma modell, se fotnot 25). Modell 6 visar att en ytterligare stjärna ger omedelbar effekt på inflödet medan en förlust inte omedelbart översätts i ett minskat inflöde. Men teoretiskt sett kan detta bero på att sådana förändringar tar längre tid för deltagare att reagera på. Denna möjlighet undersöks vidare på sida 30.

Tabell A8 Effekt av rating på (logaritmerat) inflöde av deltagare (juni 2017–september 2019)

	(1)	(2)	(3)	(4)	(5)	(6)
Typ av transformering eller begränsning på ratingen:	Harmoniserad rating, exkl. utan rating	Harmoniserad rating	Rating, exkl. utan rating	Rating	Rating	Positiv/negativ förändring av rating
Period:	Hela perioden	Hela perioden	Efter	Före	Efter	Hela perioden
Rating	0,24** (0,03)	0,34** (0,06)	0,20**			
Före (<dec 2018)		0,85** (0,14)				
Efter (>=dec 2018)		Ref.				
Före # Rating		-0,14* (0,06)				
Efter # Rating		Ref.				

²⁵ En kompletterande modell motsvarande modell 2 har även skattats med en flexibel specifikation (ratingdummys). Effekten av en topprating skattas då till 0,24 log-differenser lägre i föreperiod jämfört med efterperioden. Även effekten av ingen rating skattas till 0,18 lägre i föreperioden. Båda dessa skillnader är statistiskt signifikanta på en 5-procentsnivå. För övriga ratingnivåer finner vi inga statistiskt säkerställda skillnader. Det är dock viktigt att notera att även om skillnaden är statistiskt signifikant kan vi inte vara säkra på att det var utökningen från 3 till 4 stjärnor som orsakade denna ökning eftersom vi inte har något kontrafaktiskt utfall att jämföra med.

²⁶ För att beräkna procentuella effekter från rating 3 till 4 subtraheras koefficienten för 3 från 4 och räknas om till procent enligt: $0,52 - 0,27 = 0,25$ log-diff = $100 * (\exp(0,25) - 1) = 28$ procent.

	(1)	(2)	(3)	(4)	(5)	(6)
Typ av transformering eller begränsning på ratingen:	Harmoniserad rating, exkl. utan rating	Harmoniserad rating	Rating, exkl. utan rating	Rating	Rating	Positiv/negativ förändring av rating
Inga stjärnor				0,08 (0,06)	0,11 (0,08)	
1				Ref.	Ref.	
2				0,08 (0,05)	0,06 (0,07)	
3				0,44** (0,06)	0,27** (0,08)	
4					0,52** (0,10)	
Oförändrad						Ref.
Negativ						0,01 (0,04)
Positiv						0,17** (0,04)
Konstant	1,51** (0,08)	0,74** (0,13)	1,15** (0,09)	1,55** (0,05)	1,22** (0,07)	1,70** (0,03)
Fixa tidseffekter	Ja	Ja	Ja	Ja	Ja	Ja
Observationer	8 218	8 218	2 311	9 256	3 626	12 882
R ²	0,209	0,210	0,186	0,147	0,124	0,140
Modell	fe	fe	fe	fe	fe	fe

Anm.: Standardfel inom parenteser är klustrade på avtal. * $p < 0,05$; ** $p < 0,01$. Modell: fe betyder fixa effekter för avtal.

Källor: Arbetsförmedlingen och egna beräkningar.

Modell 1 till 3 i tabell A9 visar ratingens påverkan på inflödet av förstagångsdeltagare där effekten skattas omedelbart (F0), efter en månad (F1) och efter två månader (F2). Det visar sig att effekten är starkare efter 1 till 2 månader efter att den nya ratingen

erhållits.²⁷ I modell 4 till 6 avses effekten av en förlorad eller vunnen rating. En leverantör som just förlorat en stjärna upplever inte ett minskat flöde av förstagångsväljare förrän ungefär 2 månader senare (effekten är dock endast marginellt signifikant). En ökad stjärna innebär dock en omedelbar effekt, även om den fördröjda effekten är något större.

Tabell A9 Fördröjd effekt av rating på (logaritmerat) inflöde av deltagare

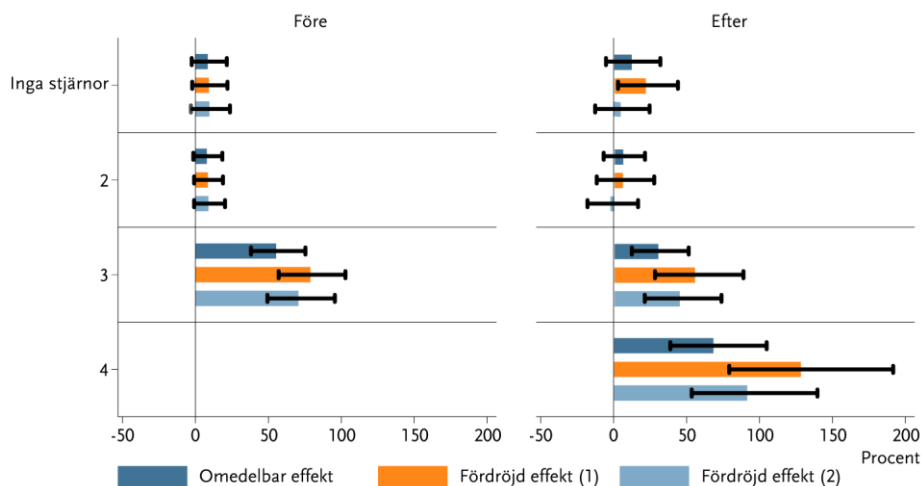
	(1)	(2)	(3)	(4)	(5)	(6)
	F0	F1	F2	F0	F1	F2
Rating harmoniserad	0,24**	0,32**	0,31**			
	(0,03)	(0,04)	(0,04)			
Oförändrad				Ref.	Ref.	Ref.
Negativ				0,01	-0,05	-0,08
				(0,04)	(0,04)	(0,04)
Positiv				0,17**	0,22**	0,21**
				(0,04)	(0,04)	(0,04)
Konstant	1,51**	1,34**	0,89**	1,70**	1,70**	1,30**
	(0,08)	(0,09)	(0,09)	(0,03)	(0,03)	(0,03)
Fixa tidseffekter	Ja	Ja	Ja	Ja	Ja	Ja
Observationer	8 218	7 743	7 281	12 882	12 297	11 611
R ²	0,209	0,227	0,230	0,140	0,152	0,159
Modell	fe	fe	fe	fe	fe	fe

Anm.: Period: juni 2017–september 2019. Standardfel inom parenteser är klustrade på avtal. * $p < 0,05$; ** $p < 0,01$.

Källor: Arbetsförmedlingen och egna beräkningar.

I figur A19 redovisas motsvarande (modell 1 till 3) men med en flexibel funktionsform och uppdelat på före- och efterperioden. Båda perioderna uppvisar samma mönster som tidigare, det vill säga effekten av en förändrad rating blir större efter 1 till 2 månader.

²⁷ Det går även att skriva om ekvationerna så att ratingen *laggas* (förskjuts bakåt) i stället för att utfallet förskjuts framåt. Då går det även att inkludera fler laggar av ratingen i en och samma ekvation. Vid en sådan specifikation kan varje lagg av ratingen tolkas som hur mycket just den månaden påverkade utfallet (till skillnad från den redovisade specifikationen då varje förskjuten effekt även inkluderar effekten innan). Men tyvärr blir de skattningarna mycket oprecisa och därmed svårtolkade. En övergripande slutsats från den analysen är dock att den största delen av den samlade effekten kommer i den första laggen, dvs. ratingen för en månad sedan.

Figur A19 Fördröjd effekt av rating på (logaritmerat) inflöde av förstagångsdeltagare

Anm.: Separata regressioner har skattats för båda perioderna samt för de olika förskjutningarna av utfallen. Period: juni 2017–september 2019. De horisontella strecken som utgår från punktestimeringarna är 95-procentiga konfidensintervall beräknade från klustrade standardfel utifrån avtal. Modell: fixa effekter för avtal.

Källor: Arbetsförmedlingen och egna beräkningar.

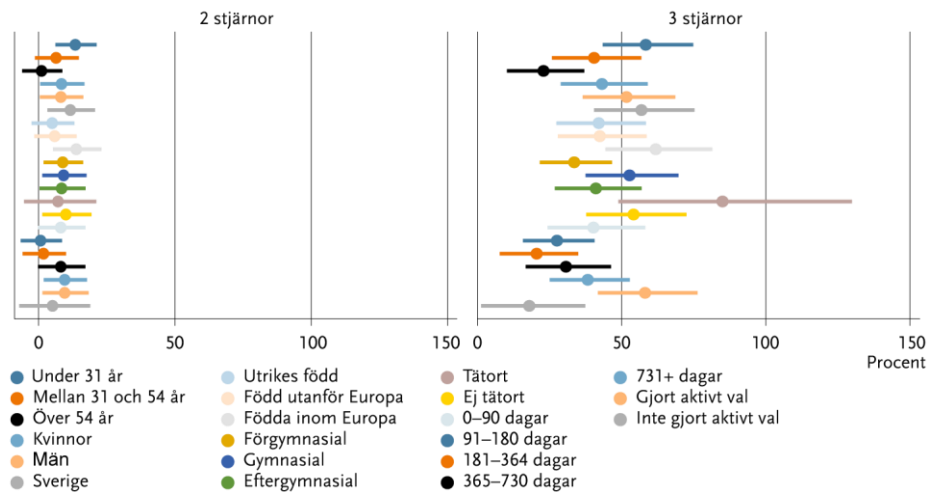
Figur A20 visar separata effektskattningar av förändrad harmoniserad rating (2 stjärnor och 3 stjärnor) för olika typer av flöden grupper av förstagångsdeltagare. Effekten av 2 stjärnor varierar inte särskilt mycket mellan grupperna. För de allra flesta är effekten noll till svagt positiv. För 3 stjärnor är skillnaderna tydligare. Exempelvis verkar yngre reagera betydligt mer på leverantörer med fler stjärnor än äldre.²⁸ Ett annat exempel är att personer som inte gör aktiva förstagångsval inte heller hamnar hos högrankade leverantörer. Men detta är skattat över hela perioden. Sedan augusti 2018 är ickevalsalternativet den högst rankade leverantören inom 20 mil från den arbetssökande, innan dess var det endast närhetsprincipen som tillämpades. Det var dock först i november 2018 som förändringen lades in i handläggarsstödet.²⁹ Figur A21 visar hur effektskattningen förändras över tid för deltagare som gjorde ett aktivt respektive passivt val. Effektstorleken ökade från och med november 2018 för deltagare med passiva val. Kort efter att regeln för hur icke-valet hanteras förändrades även antalet stjärnor i systemet. Vi vet därför inte med säkerhet om effekten hade ökat även

²⁸ För att uttala oss om skillnaderna i effektskattningar för 3 stjärnor är statistiskt signifikanta eller ej har vi genomfört formella test mellan parameterskattningarna. Dessa kommenteras här kortfattat utifrån om det finns statistiskt signifikanta skillnader, dvs. om p-värdet understiger 0,05. Skattningen för gruppen under 31 år är exempelvis inte statistiskt signifikant skild från gruppen mellan 31 och 54 år men skillnaden mot gruppen som är över 54 år är statistiskt signifikant. Skillnaden mellan inrikes och utrikes födda är inte signifikant, men däremot skillnaden mellan födda utanför Europa jämfört med födda inom Europa (där även svenskfödda ingår). Skillnaden mellan gruppen med högst gymnasieutbildning och förgymnasial utbildning är signifikant men skillnaden mellan eftergymnasialt utbildade och de förgymnasialt utbildade är inte signifikant. Vi finner inga statistiskt signifikanta skillnader mellan de som bor i tätort och övriga. Gruppen med kortast väntetid, 0–90 dagar, är statistiskt signifikant skild från gruppen med 181–364 dagars väntetid men inte gentemot övriga. Den tydligaste skillnaden finner vi mellan gruppen som gjorde ett aktivt val och de som inte gjorde det.

²⁹ Arbetsförmedlingen, *Handläggarsstöd, Kundval – Stöd och matchning*, 2018.

utan nivåutökningen men utan regelförändringen av hur passiva val hanteras hade nivåutökningen inte kunnat påverka gruppen alls.

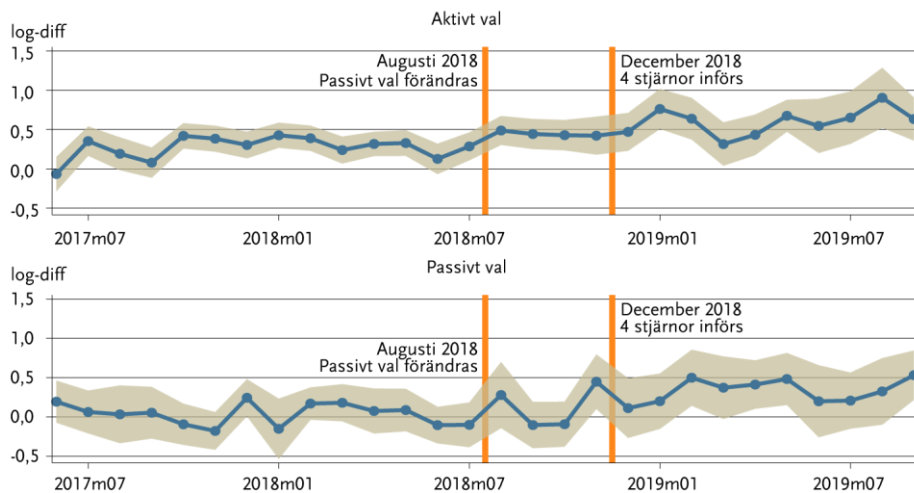
Figur A20 Heterogena effekter



Anm.: Period: juni 2017–september 2019. De horisontella strecken som utgår från punktestimeringarna är 95-procentiga konfidensintervall beräknade från klustrade standardfel utifrån avtal.

Källor: Arbetsförmedlingen och egna beräkningar.

Figur A21 Effekt av topprating fördelat på aktivt/passivt val



Anm.: Effekten mäts i log-differenser. Topprankad avser 3 stjärnor innan december 2018 och sedan 4 stjärnor. Referensgruppen innehåller de andra rating-nivåerna (exkl. avsaknad av rating). 95-procentigt konfidensintervall i grått fält beräknat med klustrade standardfel på avtal.

Källor: Arbetsförmedlingen och egna beräkningar.

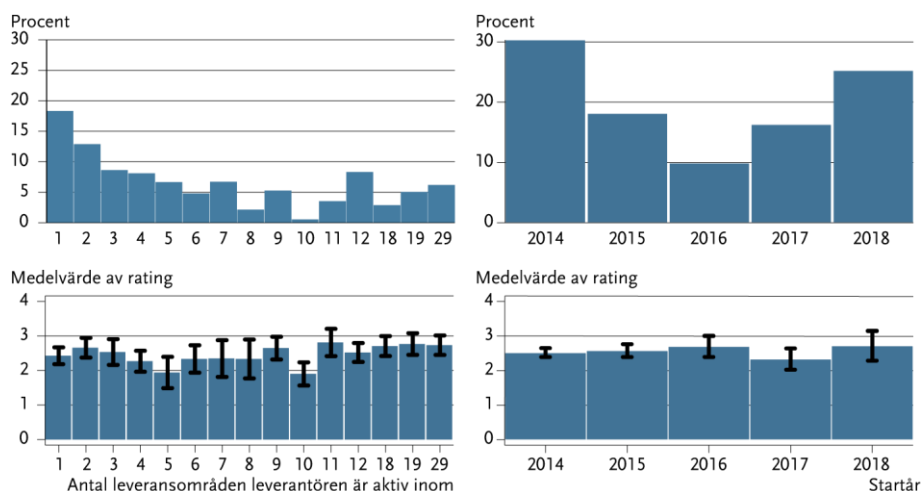
5.3 Samvariationen mellan rating och leverantörens storlek och tid i systemet

Analysen som följer visar att varken leverantörernas storlek eller tid i systemet har något betydande förklaringsvärde på ratingen. Det tyder på att även relativt nya leverantörer kan etablera sig och uppnå konkurrenskraftiga resultat.

Stora leverantörer definieras i rapporten som leverantörer som etablerar sig i flera leveransområden, det vill säga de har flera avtal ute i landet. Vi har valt denna definition eftersom den typen av information finns tillgänglig med hög aktualitet, till skillnad från företagsekonomiska data som släpar efter i flera år. Tid i systemet mäts med dummys för leverantörens startår i STOM.

Figur A22 visar histogram över leverantörsstorlek och startår för leverantörerna samt medelvärden av ratingen fördelat på samma variabler. Här framgår inga tydliga mönster för någon analysvariabel. En kompletterande regressionsanalys har även genomförts för både leverantörsstorlekens samvariation på ratingen och leverantörens startår. Här framkommer ett svagt positivt linjärt samband för leverantörsstorlek och rating. Men lutningskoefficienten är mycket låg och förklaringsvärdet därmed försumbart. Det predicerade rating-medelvärdet för en leverantör som endast finns etablerad i ett leveransområde är cirka 2,5 medan det för de allra största leverantörerna är cirka 2,65. När vi grupperar leverantörerna från små till stora leverantörer finner vi dessutom ingen statistiskt signifikant skillnad mellan de minsta (1 område) och de största (fler än 10 områden). Inte heller startåret för avtalet visar på något samband med ratingen.

Figur A22 Samband mellan rating och leverantörsstorlek eller startår för avtalet (december 2018–september 2019)



Källa: Arbetsförmedlingen.

5.4 Kvalitet i R2 mellan leverantörerna

Från analysen i avsnitt 5.2 framkommer att ratingen påverkar både flödet av STOM-deltagare och överlevnadsgraden för leverantörerna. Därför är det intressant att undersöka kvaliteten i utfallsmåttet R2 som Arbetsförmedlingen använder sig av. Utfallet R2 innebär att den arbetssökande i STOM har fått ett, av Arbetsförmedlingen, godkänt utfall. Ett godkänt utfall innebär att en anställning som motsvarar hela den arbetssökandes arbetsförmåga³⁰ har påbörjats senast 2 månader efter avslutat deltagande i STOM hos en given leverantör och att anställningen har pågått i minst 4 månader. Ett godkänt utfall kan även vara att den arbetssökande har påbörjat en utbildning. Kvaliteten i R2 undersöks dels kortsiktigt genom månadsuppgifterna från Skatteverket, dels på lite längre sikt (omkring ett år) genom registeruppgifter från LISA (exempelvis kontrolluppgifter som sammanställts på helår).

5.4.1 Månadsuppgifterna och STOM

När en deltagare får ett jobb eller påbörjar utbildning registreras ett så kallat R1. Den arbetssökande avregistrerar sig även från Arbetsförmedlingen. Godkänd R2 är baserad dels på inrapportering från leverantören, dels på att deltagaren inte åter registrerar sig hos Arbetsförmedlingen.

Skatteverkets månadsuppgifter möjliggör kontroll av hur R2 korresponderar med registeruppgifterna. Från och med 2019 ska alla utbetalda löner registreras hos Skatteverket varje månad. I teorin ska alltså en deltagare som fått R1 i januari 2019 fått minst fyra utbetalda löner fem månader senare³¹ från den angivna arbetsgivaren.

Det fanns 4 737 deltagare med godkänd R1 och R2 under 2019. Av dessa återfanns 4 376 hos Skatteverket.³² Av dessa R2 var 331 av typen utbildning och exkluderas således från denna analys. Analyspopulationen består således av 4 045 individer fördelat på 36 907 observationer.

Enligt reglerna Arbetsförmedlingen satt upp ska en godkänd R2 för arbete föregås av minst 4 månaders anställning från och med R1 (då anställning först rapporterades in). För att undersöka detta ställer vi upp olika inkluderingskriterier. Exempelvis räknar vi antal månader med en registrerad arbetsgivare eller om de får inkomster över vissa gränser under resultatperioden. Vi har utökat uppföljningsfönstret med en månad eftersom lön ofta betalas ut i efterskott. Denna period är oftast 4 + 1 månad men ibland längre. För enkelhetens skull kommer vi härnäst referera till denna period som *resultatperioden*.

Av den totala analyspopulationen på 4 045 individer hade 4 041 minst en registrerad arbetsgivare under resultatperioden. Av dessa fick 90 procent (3 631 personer) minst en månad med mer än 16 000 kr i lön. Nästan 4 000 (99 procent) individer hade en registrerad arbetsgivare i fyra månader under resultatperioden.

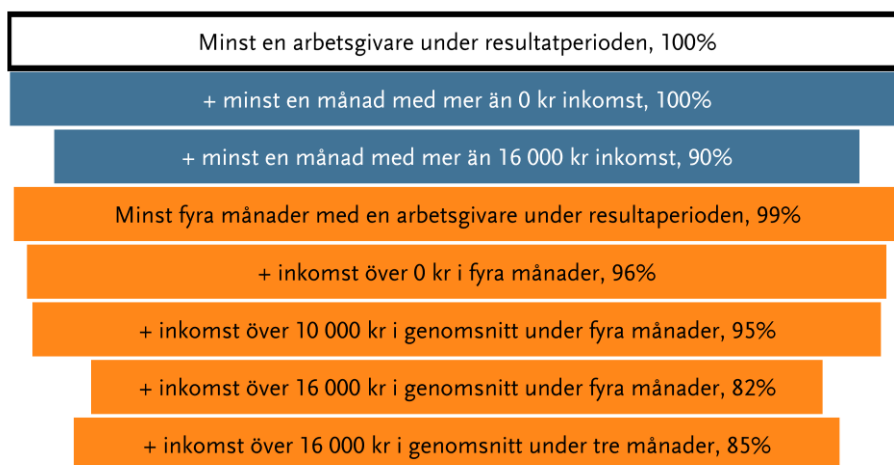
³⁰ Om den arbetssökandes arbetsförmåga bedöms till exempelvis 50 procent krävs således en anställning på minst 50 procent.

³¹ Fyra månader + en för eventuellt släpande löneutbetalning. I vissa fall överskrider dock tiden mellan R1 och R2 fyra månader. Detta tar vi individuell hänsyn till. Vi kallar denna period för resultatperioden.

³² Bortfall pga. exempelvis skyddade identiteter.

Drygt 3 870 (96 procent) hade inkomster över 0 kr för minst fyra månader, och cirka 90 procent hade en medelinkomst som översteg 10 000 kronor under de fyra bäst betalda månaderna i resultatperioden.³³ Om vi höjer gränsen till 16 000 kronor i månaden blir andelen som klarar gränsen betydligt lägre, 82 procent.³⁴ Det kan finnas flera orsaker till att så många hamnar under gränsen när den sätts vid 16 000. Om en deltagares arbetsutbud inte är 100 procent krävs heller inte att deras anställning ska omfatta mer än deras arbetsutbud. Dessutom kan halv resultatersättning betalas ut om anställningen inte motsvarar hela, men mer än 50 procent av, deltagarens önskade arbetsutbud. Medianlönen i gruppen ligger omkring 20 000 kronor, vilket innebär att en halvtid ger 10 000 kronor. Om vi analyserar lönedistributionen framkommer även att det finns en tydlig ansamling precis vid 10 000 kronor (se figur A24). Men det är även möjligt att en viss andel av dessa anställningar utbetalar en månadslön som är betydligt lägre än de kollektivavtalade lägsta lönerna. För att kunna slå fast något sådant krävs dock en fördjupad undersökning av de enskilda fallen.

Figur A23 Antal individer med godkända R2 som återfinns hos Skatteverket enligt olika kriterier

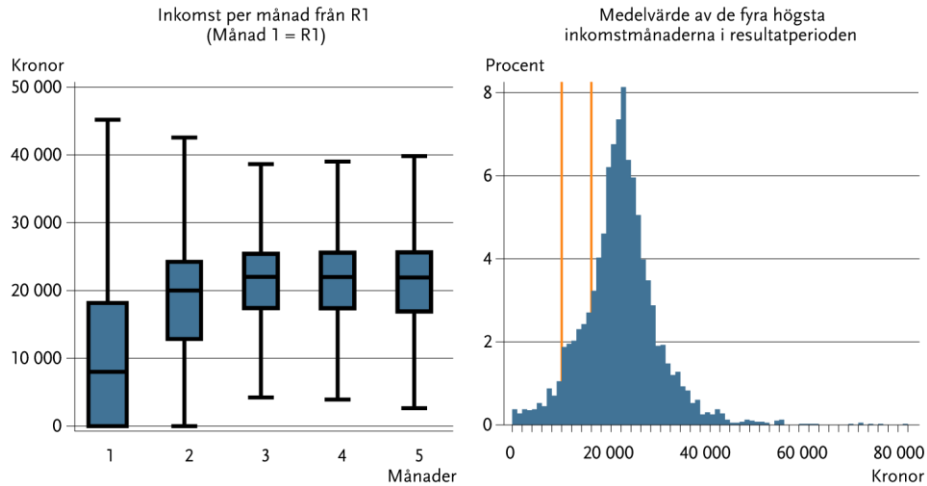


Källor: Arbetsförmedlingen och Skatteverket.

Av de individer som återfinns inom resultatperioden är medianinkomsten cirka 20 000 kr till 22 000 kr för månad 2 till 5 (se figur A24). Månad 1 är betydligt lägre eftersom lön ofta betalas ut i efterskott. För varje månad i resultatperioden finns det anmärkningsvärda maximalbelopp som betalats ut på mellan cirka 100 000 och 180 000 kr.

³³ Nivån på de registrerade månadsinkomsterna är inte alltid ett bra mått på om det finns ett anställningskontrakt med tillräcklig månadslön. Vissa arbetsgivare betalar ut lön i efterskott och vissa i förskott. Vissa arbetsgivare betalar ut semesterersättning och vissa erbjuder ledighet med lön och semester tillägg. Det kan även finnas internadministrativa tillfälliga orsaker till att lönen inte alltid inkommer i ett jämnt flöde. För att i möjligaste mån beakta dessa förklarliga variationer när vi beräknar genomsnittslönen väljer vi ut de tre eller fyra månader som har högst registrerade inkomster för individerna under resultatperioden (som oftast är fem månader).

³⁴ Andelen ökar till 85 procent om vi begränsar beräkningen till de tre bäst betalda månaderna.

Figur A24 Inkomstspridning under resultatperioden

Anm.: De två vertikala strecken i histogrammet markerar 10 000 och 16 000 kronor. Resultatperiod avser perioden mellan R1 och R2 + 1 månad. Förklaring till låddiagrammet: Lådan omfattar den 1:a till den 3:e kvartilen. Strecket i lådan markerar medianen. Slutet på de rätvinkliga strecken som utgår från lådan är maximum- respektive minimumvärdet rensat för extremvärden. Dessa gränser beräknas genom att multiplicera den 1:a respektive 3:e kvartilen med 1,5.

Källor: Arbetsförmedlingen, Skatteverket och egna beräkningar.

Sammanfattningsvis tycks de inrapporterade uppgifterna om en godkänd R2 korrespondera relativt väl med Skatteverkets månadsuppgifter. Men att cirka 5 procent inte hade mer än 10 000 kronor i genomsnittslön i resultatperioden tyder på att en viss andel av dessa anställningar kanske avviker kraftigt från kollektivavtalsenliga villkor på svensk arbetsmarknad eller är av en lägre omfattning än vad som krävs för egentligt godkänt R2.

5.4.2 Helårsuppgifter från LISA och STOM

För att tillse en tillräcklig mängd observationer samtidigt som uppföljningstiden blir tillräckligt lång har vi valt ut de individer som fått godkänt R2 mellan juni 2016 och januari 2017. Dessutom ska deras leverantör ha haft en rating.

Tabell A10 redovisar ett antal arbetsmarknadsutfall som inte borde skilja sig åt mellan de olika ratingarna. Modell 1 visar att de arbetssökande i STOM som fick godkänt R2 genom en leverantör med 1 stjärna hade 73 procents sannolikhet att vara registrerad som förvärvsarbetande i november 2017 (RAMS-definitionen). Vi finner inga statistiskt säkerställda skillnader gentemot individer som erhållit godkänt R2 från leverantörer med högre rating. Vi finner heller inga statistiskt säkerställda skillnader för något av de andra utfallen. Samtliga modeller har en mycket låg förklaringsgrad, vilket innebär att uppdelningen på stjärnor inte är ett meningsfullt sätt att förstå skillnaderna i utfall för individerna. Ur ett systemperspektiv är detta positivt eftersom kvaliteten i R2 därmed kan betraktas som likvärdig. Huruvida 73 procents förvärvsintensitet i gruppen är att betrakta som tillräckligt bra kan vi dock inte säga något om i den här typen av analys.

Tabell A10 Regressionstabell över kvalitet i R2 – Individer som fick godkänt R2 mellan juni 2016 och januari 2017 mot utfall för 2017

	(1)	(2)	(3)	(4)	(5)
	Sysselsatt enligt RAMS	Har haft inkomster under året men inte tillräckligt för att uppfylla RAMS-sysselsatt	Förvärvsinkomst (tkr)	Kvantilregression (median, Förvärvsinkomst [tkr])	ln(Förvärvsinkomst)
1 stjärna	Ref.	Ref.	Ref.	Ref.	Ref.
2 stjärnor	0,01 (0,02)	-0,01 (0,01)	7,04 (310,23)	3,50 (245,93)	0,04 (0,04)
3 stjärnor	-0,01 (0,02)	0,01 (0,02)	-1,78 (86,97)	-8,00 (626,96)	-0,03 (0,04)
Konstant	0,73** (0,01)	0,19** (0,01)	186,28** (7565,78)	197,10** (12837,91)	7,30** (0,04)
Obs.	8 755	8 755	8 712	8 712	8 003
R ²	0,000	0,000	0,001		0,001

Anm.: Regressionen är skattad för individer som fick godkänt R2 mellan juni 2016 och januari 2017. Robusta standardfel. * $p < 0,05$; ** $p < 0,01$.

Källor: Arbetsförmedlingen, Statistiska centralbyrån och egna beräkningar.

Olika leverantörer kan dock fokusera olika mycket på att få deltagare till jobb eller utbildning. Om exempelvis de lägst rankade leverantörerna har mindre fokus på utbildning för att nå godkänt R2 riskerar vi att göra orättvisa jämförelser mellan grupperna. Riksrevisionen har dock bara tillgång till fördelningen mellan jobb och utbildning på leverantörsnivå (inte på individnivå). Därför måste en sådan analys aggregeras upp till leverantörsnivå genom att beräkna viktade medelvärden av utfallen (såsom förvärvsinkomster). Därefter kan vi inkludera en kontroll för *andelen till jobb i R2*. En sådan analys visar dock att skillnaderna mellan ratingnivåerna fortsatt var statistiskt insignifikanta med och utan kontrollen för andel till arbete i R2.³⁵

³⁵ För den undersökta perioden bland leverantörer med en fastställd rating var andelen till arbete av R2 cirka 87 procent. Det fanns här statistiskt säkerställda skillnader mellan grupperna (5 procent fler till arbete för rating 2 och 4 procent mer för rating 3 i jämförelse med rating 1). Även om skillnaden korrelerade signifikant med exempelvis förvärvsinkomst påverkade det ändå knappt skillnaden mellan leverantörerna i slutändan.

Referenser

Allison, P.D., "Event History and Survival Analysis: Introduction", *Event History and Survival Analysis*, 2014.

Arbetsförmedlingen, *Beskrivning av hur ratingen tas fram i tjänsten Arbetsförmedlingens kundval – Stöd och matchning*, Enheten Forskning och Utvärdering (hämtad 2020-01-28) Arbetsförmedlingen.

Arbetsförmedlingen, *Handläggarsöd, Kundval – Stöd och matchning*, version 4.0, beslutsdatum 2018-11-23, Arbetsförmedlingen, 2018.

Arbetsförmedlingen, *Kundval - Stöd och matchning*, Arbetsförmedlingen, 2018.

Gerdes, C., "Does Performance Information Affect Job Seekers in Selecting Private Providers in Voucher-Based ALMP Programs? | IZA - Institute of Labor Economics", *IZA DP*, nr 8992, 2015.

Jenkins, S.P., "Easy Estimation Methods for Discrete-Time Duration Models", *Oxford Bulletin of Economics and Statistics*, vol. 57, nr 1, 1995.

SOU 2019:3, *Effektivt, tydligt och träffsäkert – det statliga åtagandet för framtidens arbetsmarknad*, 2019.