

Vad avgör träffsäkerheten i bedömningar av arbetssökandes stödbehov?

En undersökning av förutsättningarna för statistiska bedömningar av avstånd till arbetsmarknaden, med fokus på betydelsen av inskrivningstid

© Arbetsförmedlingen
Författare: Petter Helgesson och Petra Ornstein
Datum: 2021-04-12
Diarienummer: Af-2020/0046 8022

Innehåll

1	Sammanfattning	5
1.1	Statistiska verktyg för att möjliggöra träffsäker tilldelning av insatser	5
1.2	Jämförelse av träffsäkerhet för nyinskrivna och befintliga sökande ("sökandestocken")	6
1.3	Bättre träffsäkerhet för sökandestocken.....	6
1.4	Fler undersökta frågor	6
1.5	Rekommendationer	7
2	Inledning	8
2.1	Bakgrund	8
2.1.1	Statistiskt bedömningsstöd och ökad träffsäkerhet.....	8
2.1.2	Viktiga frågeställningar i denna rapport.....	8
2.2	Rapportens disposition.....	10
3	En statistisk modell för bedömning av avstånd till arbetsmarknaden	11
3.1	Modellerna bygger på inskrivna under 2018	12
3.2	Avaktualiserade med okänd orsak och inskrivningar från före 2012 rensas bort	13
4	Träffsäkerheten definieras som en viktad andel som prediceras "rätt"	14
4.1	Definition och motivering av träffsäkerhetsmättet	14
4.2	Hur träffsäkerheten beräknas – steg för steg	16
5	Resultat	17
5.1	Träffsäkerheten är relativt god vilket gör undersökningen relevant	17
5.2	Skillnaden i träffsäkerhet mellan sökandestocken och nyinskrivna är stor	18
5.3	Information ökar och uppdateras under inskrivningen vilket kan förklara skillnaden mellan nyinskrivna och sökandestock.....	19
5.4	Träffsäkerheten påverkas mer av testpopulationen än av träningspopulationen	20
5.5	Träffsäkerheten är inte allt	22
5.6	Uppdaterade sökandekategorier viktigaste informationstillskottet – följt av inskrivningstid	23
6	Sammanfattning av undersökningen och dess resultat	24
6.1	Ett nytt träffsäkerhetsmått tas fram för att kunna jämföra träffsäkerhet mellan olika populationer	25
6.2	Högst träffsäkerhet fås för sökandestocken.....	25
6.3	Sökandestockens högre träffsäkerhet beror till stor del på tillkommande information under tiden som inskriven.....	25
6.4	Skillnader i träffsäkerhet beror främst på vilken population modellen tillämpas på, oavsett vilken population modellen tränats på	25
6.5	Modell för nyinskrivna har hyfsad träffsäkerhet på sökandestocken.....	26

6.6	Modell tränad på ett visst år går bra att använda på ett annat år.....	26
7	Slutsatser och rekommendationer	26
7.1	Tillkommande information ger högre träffsäkerhet – överväg att avvakta med dyrare insatser och säkerställ datainsamling under inskrivningstiden	26
7.2	Undersök gärna möjligheten till ökad träffsäkerhet, men överskatta den inte	27
7.3	Studera hur möjligheten till träffsäkerhet beror på frågan som ställs.....	27
7.4	Träffsäkerhet är inte allt – studera även systematiska fel när modeller används för en annan population	28
7.5	Modell från studerade år går att använda på andra år – studera detta under mer dramatiska förändringar	28
	Referenser	28
	Bilaga 1. Lite mer om modellen	29
	Bilaga 2. Viktningen i träffsäkerhetsmåttet	34
	Bilaga 3. Populationernas storlek.....	35

1 Sammanfattning

En viktig del av Arbetsförmedlingens reformering är en ökad grad av digitalisering och automatisering av verksamheten. Som del av denna process utvecklar Arbetsförmedlingen kunskapen om förutsättningar och möjligheter för att använda statistiska verktyg i myndighetsutövningen. Denna rapport är en del i denna kunskapsutveckling. En ökad förmåga att nyttja statistiska bedömningsverktyg har potential att bidra starkt till förutsättningarna för att automatisera delar av handlägningsprocessen och ingår som en viktig del i reformeringsarbetet.

Den typ av statistiska verktyg som studeras i denna rapport syftar till att bedöma arbetssökandes avstånd till arbetsmarknaden och därmed deras stödbehov. Hos ett sådant verktyg blir det centralt med god träffsäkerhet: att korrekt kunna förutsäga hur lång tid det tar för en arbetssökande hos Arbetsförmedlingen att hitta ett jobb. Denna rapport undersöker hur träffsäkerheten kan ökas genom olika sätt att utforma och använda ett sådant verktyg, med fokus på en jämförelse mellan träffsäkerhet för nyinskrivna och befintliga sökande ("sökandestocken"). En annan fråga som undersöks är hur träffsäkerheten påverkas av att ett statistiskt verktyg "flyttas" mellan olika populationer. Rapporten berör också vilken typ av träffsäkerhet som kan förväntas från ett statistiskt verktyg som bedömer arbetssökandes stödbehov.

Rapportens huvudsakliga slutsatser redovisas i sammanfattningen, som också avslutas med några rekommendationer till hur Arbetsförmedlingen kan fortsätta arbetet med att förbättra träffsäkerhet och rättssäkerhet i myndighetens arbetsmarknadspolitiska bedömningar. För mer ingående reflektioner kring rapportens resultat och slutsatser hänvisas till avsnitt 6 och 7.

1.1 Statistiska verktyg för att möjliggöra träffsäker tilldelning av insatser

En träffsäker fördelning av arbetsmarknadspolitiska insatser förutsätter att Arbetsförmedlingen kan bedöma arbetssökandes förutsättningar på arbetsmarknaden. Dessa förutsättningar påverkar nämligen hur mycket stöd den arbetssökande kan förväntas behöva för att hitta ett arbete.

Det är väl känt att arbetssökande med en längre pågående tid utan arbete har en större risk för fortsatt lång arbetslöshet. Därför har arbetsmarknadspolitiska insatser historiskt sett främst tilldelats arbetssökande efter en längre inskrivningstid på Arbetsförmedlingen. Under det senaste decenniet har dock initiativ tagits för att tilldela insatser tidigare under arbetslöshetstiden till arbetssökande med större risk för långtidsarbetslöshet. För att möjliggöra detta kan ett statistiskt verktyg användas som stöd: det finns starkt stöd i forskningen för att sådana verktyg kan förväntas öka träffsäkerheten i bedömningar av stödbehov. I vilken grad träffsäkerheten begränsas av att bedömningen sker tidigt under individens arbetslöshetstid är dock mindre

känt. Det kommer bero på vilken roll inskrivningstiden spelar för att identifiera arbetssökande med stora stödbehov, och det är temat för denna undersökning.

1.2 Jämförelse av träffsäkerhet för nyinskrivna och befintliga sökande ("sökandestocken")

Denna rapport undersöker hur träffsäkerheten i en statistisk modell som bedömer arbetssökandes avstånd till arbetsmarknaden skiljer sig mellan nyinskrivna och sökandestocken. Sökandestocken inkluderar alla som är inskrivna hos Arbetsförmedlingen vid en viss tidpunkt, oavsett hur länge de har varit inskrivna. Till skillnad från hos de nyinskrivna varierar alltså inskrivningstiden. För sökandestocken kan denna variation utnyttjas som en del i bedömningen vilket alltså inte är möjligt för de nyinskrivna. Därutöver tillkommer information under tiden som inskriven på Arbetsförmedlingen, vilken också kan utnyttjas när den förväntade (kvarstående) arbetslöshetstiden ska bedömas.

1.3 Bättre träffsäkerhet för sökandestocken

När träffsäkerheten jämförs mellan nyinskrivna och sökandestocken visar resultaten på en tydligt bättre träffsäkerhet för sökandestocken. Möjligheten att utnyttja inskrivningstiden i bedömningen har en viss betydelse för detta. Det visar sig dock att annan information som tillkommer eller uppdateras under tiden som inskriven är ännu viktigare. Allra viktigast är uppdateringar av de arbetssökandes så kallade sökandekategorier, som bär med sig viktig information om händelser under tiden som arbetslös, till exempel deltagande i olika insatser.

1.4 Fler undersökta frågor

Ytterligare ett antal besläktade frågor undersöks och diskuteras i rapporten. Bland annat framgår att modeller i många fall kan byggas på en grupp arbetssökande och sedan användas på en helt annan grupp arbetssökande, med delvis skilda förutsättningar och villkor, utan att träffsäkerheten påverkas särskilt mycket. Till exempel kan en modell byggd på data från ett år användas för andra år. Det här är ett viktigt resultat eftersom modeller av det här slaget alltid måste byggas på historiska data. Det går även att bygga en modell på nyinskrivna och använda den på sökandestocken med någorlunda god träffsäkerhet. Dessa observationer kan ses som exempel på en mer generell tendens för denna typ av bedömningar: detaljer i modellens utformning har en begränsad påverkan på träffsäkerheten. Ett annat exempel på detta är att den relativt enkla modell som används i denna rapport har liknande träffsäkerhet som mer komplicerade modeller.

Det ska dock poängteras att träffsäkerheten inte är det enda måttet på hur bra en modell är. En annan viktig aspekt är att modellen inte innehåller systematiska fel. Genom att till exempel använda en modell för nyinskrivna på sökandestocken uppstår en risk för systematiska fel. Exempelvis skulle långtidsarbetslösa

svenskfödda män med gymnasieutbildning få stödbehovet systematiskt underskattat om inte deras arbetslöshetstid togs med i bedömningen. Risk för systematiska fel föreligger även vid stora förändringar över tid. Hur risken för systematiska fel i bedömningar av stödbehov påverkas vid stora variationer över tid eller externa chocker som påverkar ekonomin och arbetsmarknaden behöver undersökas närmare.

1.5 Rekommendationer

Utifrån den här rapporten kan följande rekommendationer lämnas:

- Vid anvisningar till insatser och i utformning av ersättningssystem: ha i åtanke att bedömningar av avstånd till arbetsmarknaden alltid är behäftade med osäkerhet. Möjligheten att öka träffsäkerheten genom modellutveckling och tillgång till ytterligare data är begränsad.¹ Detta har policyimplikationer. Exempelvis behöver kontrollsystem och marknadsmekanismer på en framtida leverantörsmarknad hantera att stödbehovet kommer att underskattas för vissa arbetssökande och överskattas för andra.
- Beakta att olika frågor är olika svåra att svara på. Exempelvis är det lätt att nå hög träffsäkerhet om målet är att skilja ut de arbetssökande som står allra längst ifrån arbetsmarknaden från de som står allra närmast. Det är betydligt svårare att skilja arbetssökande i olika ”mellanskikt” från varandra. Sådan uppdelning i mellanskikt bör därför användas med försiktighet, exempelvis i utformningen av ersättningssystem.
- Det kan vara värt att vänta en tid med större satsningar och bredare tilldelning av insatser och istället skapa system för att genomföra nya bedömningar efter hand. Resultaten i denna rapport visar att det finns bättre beslutsunderlag för arbetssökande som har en viss tids pågående arbetslöshet bakom sig. Hur lång tid det är värt att vänta blir en komplicerad avvägning mellan träffsäkerheten och negativa konsekvenser av väntetid respektive inlösning och kostnader för insatser. Denna avvägning beror också på hur långt ifrån arbetsmarknaden en arbetssökande bedöms stå: tidiga insatser kan riktas till de som bedöms stå allra längst ifrån arbetsmarknaden medan mer information kan vara nödvändig i ”mellanskikten”. Avvägningen bör studeras närmare.
- Uppdaterade, korrekta och detaljerade uppgifter behöver fortsatt samlas in under pågående inskrivning, även i en reformerad myndighet. Använd denna information för att nå en god träffsäkerhet vid senare bedömningar.
- Undersök risken för systematiska fel när en modell framtagen för en viss population används på en annan population.

¹ Bättre utfallsdata i form av inkomstuppgifter från skatteverket har däremot en särställning eftersom detta skulle förhindra ett stort systematiskt bortfall, vilket sannolikt är viktigt för träffsäkerheten. Även information som samlas in under inskrivningstiden har en särskild karaktär och ett särskilt värde (se senare punkt).

2 Inledning

2.1 Bakgrund

2.1.1 Statistiskt bedömningsstöd och ökad träffsäkerhet

I Arbetsförmedlingens regleringsbrev 2020² fick myndigheten i uppdrag att utveckla ett statistiskt bedömningsstöd som ska utgöra en integrerad del av myndighetens arbetsmarknadspolitiska bedömning. Det ska i ökande utsträckning användas för att bedöma arbetssökandes stödbehov så att effektiv tilldelning av resurser möjliggörs.

En central tanke med att nyttja ett statistiskt verktyg är att verktyget ska ge god träffsäkerhet i bedömningarna: alltså att stödbehovet i stor utsträckning bedöms korrekt. Lite förenklat ska de som behöver stöd bedömas som att de behöver stöd, annars riskerar dessa arbetssökande att vara onödigt länge i arbetslöshet. Samtidigt är det önskvärt att de som *inte* behöver stöd inte ges stöd i onödan – detta riskerar att låsa in arbetssökande i insatser som dessutom utgör en kostnad för samhället.³ Att ha god träffsäkerhet i algoritmbaserade avståndsbedömningar är alltså en viktig del i att på ett bra sätt utföra uppdraget att integrera denna typ av statistiska verktyg i den arbetsmarknadspolitiska bedömningen.

Arbetsmarknadspolitiken har under lång tid dominerats av insatser till arbetssökande som varit arbetslösa en längre tid, eftersom dessa i allmänhet har en större risk för fortsatt lång arbetslöshet (Hartman 2018). Under början av 2010-talet skiftades fokus delvis mot att ge tidigare insatser till individer med hög risk för långtidsarbetslöshet (Arbetsmarknadsdepartementet 2013). Denna efterfrågan kvarstår alltjämt, inte minst i och med pågående pandemi, vilket till exempel reflekteras i Arbetsförmedlingens reviderade verksamhetsplan för 2020.⁴ För en sådan strategi är det helt centralt att kunna bedöma risken för långtidsarbetslöshet i ett tidigt skede. Statistiska verktyg har god potential att nå en relativt hög träffsäkerhet (Hartman 2018, Bennmarker, Carling & Forslund 2007), men det är mindre känt i vilken utsträckning träffsäkerheten begränsas av att bedömningen sker tidigt under inskrivningen.

2.1.2 Viktiga frågeställningar i denna rapport

Träffsäkerhet för nyinskrivna jämfört med sökandestock

Med utgångspunkt i det ovanstående fokuserar denna rapport på frågan:

² <https://www.esv.se/statsliggaren/regleringsbrev/?RBID=20264>

³ Statistiska bedömningar av en arbetssökandes avstånd till arbetsmarknaden (även kallade profilering) ska inte förväxlas med så kallad *targeting*: bedömningar av vem som ska få vilken insats. För att besvara frågan om *targeting* krävs helt andra metoder och god kunskap om vilka insatser som är mest effektiva för olika individer.

⁴ En av fyra huvudpunkter i verksamhetsplanen är ”Tidiga insatser för att förhindra långtidsarbetslöshet” (Arbetsförmedlingen 2020b).

1. Hur skiljer sig möjligheten att göra en träffsäker bedömning av en arbetssökandes avstånd till arbetsmarknaden mellan nyinskrivna och sökandestock?⁵

Som noterades i avsnitt 2.1.1 har individer som varit arbetslösa en längre tid en större risk för fortsatt arbetslöshet. Detta samband illustreras också på ett enkelt sätt i figur 1, som visar andelen arbetssökande som varit inskrivna en viss tid som gått ut i arbete eller studier ett år senare.⁶ Vid inskrivningstillfället har alla samma inskrivningstid – ingen alls – så det går då inte att använda sig av inskrivningstiden i bedömningen av hur långt ifrån arbetsmarknaden en arbetssökande står. Det finns heller inte tillgång till annan information som kommer in under inskrivningstiden.⁷ Det gör att ett verktyg utformat för nyinskrivna har andra förutsättningar än ett verktyg utformat för sökandestocken. Eftersom inskrivningstiden har ett starkt samband med avståndet till arbetsmarknaden kan skillnaden avseende dessa förutsättningar bli viktig.

Frågan undersöks genom att jämföra träffsäkerheten som uppnås med en statistisk avståndsbedömning för nyinskrivna respektive för sökandestocken. Svaret ger implikationer på hur inskrivningstid bör användas i arbetsmarknadspolitiska bedömningar. Inte minst påverkar det vilka typer av insatser som kan vara rimliga att anvisa även nyinskrivna till – vid en god träffsäkerhet även för nyinskrivna så kan kostsamma insatser med risk för inlåsnings komma i fråga även för nyinskrivna.

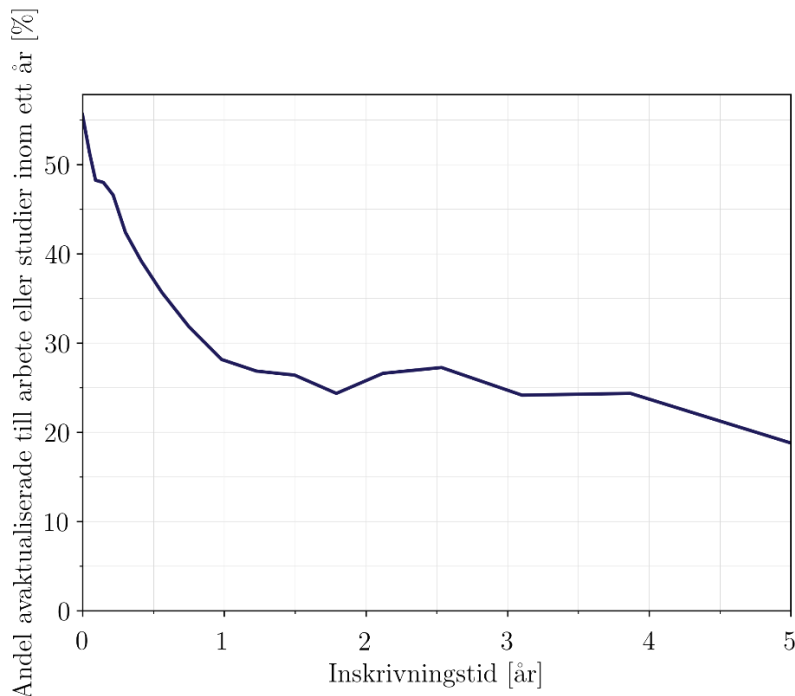
För att förstå svaret och implikationerna bättre undersöks också *vad* som avgör vad träffsäkerheten blir. Framför allt undersöks i vilken grad träffsäkerheten påverkas av *information* om inskrivningstid och annan information som blir tillgänglig under arbetslöshetstiden.

⁵ Begreppet *sökandestock* syftar på de arbetssökande som är inskrivna på Arbetsförmedlingen vid ett visst tillfälle. De *nyinskrivna* är sådana som precis har skrivits in. Mer detaljer återfinns i avsnitt 3.1.

⁶ Datat gäller individer som varit inskrivna under 2018 – ”sökandestocken” i avsnitt 3.1.

⁷ Till exempel tillkommer information om deltagande i olika insatser eller program, vilken i sig bland annat innehåller bedömningar gjorda av både arbetsförmedlare och arbetssökande.

Figur 1. Andelen som avaktualiserats till arbete eller studier ett år efter slumpvis datum som inskriven, som funktion av inskrivningstid. Populationen är densamma som sökandestocken i avsnitt 3.1 – ett urval av individer inskrivna någon gång under 2018.⁸



Hur påverkas träffsäkerheten av att byta population?

I arbetet med att öka träffsäkerheten finns en mängd andra frågor som är viktiga att undersöka närmare. Till exempel är det intressant hur en modell framtagen med hjälp av en viss population fungerar för en annan, något annorlunda, population. I praktisk användning är man alltid i den situationen i någon utsträckning: modellbygget kräver att man känner till verkliga utfall, vilket kräver en viss uppföljningstid. Samtidigt vill man använda bedömningen för att bedöma framtida utfall. För verktyget som används inom Rusta och Matcha (Krom) finns ytterligare en sådan skillnad: det är i grunden baserat på nyinskrivna och vid användning på sökandestocken hanteras inskrivningstid separat från övriga faktorer, baserat på en mer begränsad empirisk analys än övriga faktorer (Arbetsförmedlingen 2020a).

2.2 Rapportens disposition

Stora delar av denna rapport är ganska teknisk. För den läsare som vill läsa mer än sammanfattningen men vill undvika de mer tekniska delarna rekommenderas framför allt:

- Avsnitt 2.1 som ger en bakgrund och beskriver frågeställningen,

⁸ Figuren är uppbyggd med hjälp av grupper av 2 000 sökande med så näraliggande inskrivningstider som möjligt – den redovisade inskrivningstiden för varje grupp är gruppens medelvärde.

- Avsnitt 6 som innehåller en lite mer djuplodande sammanfattning av undersökningen och dess resultat,
- Avsnitt 7 som diskuterar slutsatser och rekommendationer i lite mer detalj.

De mer tekniska delarna utgörs av avsnitt 3 till 5, samt bilagorna. Avsnitt 3 beskriver översiktligt uppbyggnaden av de modeller vars träffsäkerhet senare undersöks och vilka populationer av arbetssökande de bygger på. En del detaljer i modellbyggandet återfinns i bilagorna.

Den definition av träffsäkerhet som används i texten definieras i avsnitt 4. Först diskuteras kort varför det inte är helt enkelt att jämföra träffsäkerheter mellan olika populationer. Resten av texten bygger till stor del på sådana jämförelser, så den definition som används är vald för att göra detta på ett rättvist sätt.

I avsnitt 5 följer sedan en redovisning av resultat – främst träffsäkerheten som nås för olika varianter av modellen, använd på olika populationer. Resultaten beskrivs och tolkas löpande.

3 En statistisk modell för bedömning av avstånd till arbetsmarknaden

För att snabbt få en flexibel tillgång till en statistisk arbetsmarknadspolitisk bedömning på ett sätt som möjliggör detaljerad analys av vilka variabler som har betydelse, och på vilket sätt, är en egen modell byggd för denna undersökning. Det är i grunden en samling regressionsmodeller, där olika regressioner görs för olika grupper av arbetssökande. För varje grupp korrigeras fördelningen av prediktioner så att den stämmer överens med utfallen för träningspopulationen. Som förklarande variabler används liknande individkaraktäristika som de som används i till exempel AI-Centers AMPB-modell (Arbetsförmedlingen 2020a). Tiden till avaktualisering till arbete eller studier används som målvariabel⁹. Eftersom många individer inte har hunnit nå ett positivt utfall under den tillgängliga uppföljningstiden är en stor del av utfallen så kallat censurerade: de exakta utfallen är inte kända i dessa fall men det finns däremot en känd undre gräns för utfallen – uppföljningstiden – vilken går att utnyttja. En Tobit-modell används för att hantera censureringen.

Mer detaljer kring själva modellerna återfinns i Bilaga 1. I fortsättningen av detta avsnitt beskrivs de populationer som används för att träna modellerna och för att studera deras träffsäkerhet.

⁹ I AI-Centers AMPB-modell (som används i Krom) använder en något annorlunda målvariabel: sannolikheten för att ha ett visst utfall (i Krom: avaktualisering till arbete/studier men också några sökandekategorier som räknas som arbete) vid en viss tid (i Krom: sex månader). Här används i stället tiden till utfall för att behålla något mer information, till exempel skillnaden mellan att få arbete efter 7 månader och fortfarande inte ha det efter två år. Som resultaten senare i denna promemoria indikerar så spelar dessa detaljer i modellvalet förmodligen inte särskilt stor roll eftersom olika modeller till stor del ger väldigt liknande prediktioner. Resultat i t ex Salganik m fl. (2020) som diskuteras i fotnot 24 pekar också i denna riktning.

3.1 Modellerna bygger på inskrivna under 2018¹⁰

Huvudsakligen jämförs resultat från statistiska modeller som bygger på data som ska representera följande populationer:

1. Sökandestocken: den sammansättning av inskrivna arbetssökande som finns vid ett givet tillfälle.
 - Denna population väljs slumpmässigt från de arbetssökande som varit inskrivna *någon* gång under 2018.¹¹ För att populationen ska bli representativ för sökandestocken väljs de med en sannolikhet som är proportionerlig mot tiden som inskriven under året. Som startdatum¹² används för varje individ ett slumpmässigt datum under tiden som inskriven under året.
 - För att skilja ut hur stor påverkan den information som tillkommer under inskrivningstiden studeras också en variant på sökandestocken där data från varje arbetssökandes inskrivningsdatum används.
2. Nyinskrivna.
 - Denna population utgörs av de som skrevs in under året. Som startdatum används inskrivningsdatumet. För en person som skrevs in flera gånger under 2018 används en slumpmässigt vald inskrivning.

För var och en av populationerna tas två dataset ur två icke-överlappande slumpmässiga urval på 10 procent av de arbetssökande vardera. Det ena datasetet används för att bygga modellerna (i fortsättningen kallas de "träningpopulationer") och det andra datasetet används för att bestämma träffsäkerheten ("testpopulationer"); på så vis undviks en belöning av överanpassade modeller. De som ingår i urvalet till sökandestockens träningpopulation ingår också i urvalet till de nyinskrivnas träningpopulation; motsvarande gäller också för testpopulationerna.

En stor mängd data finns sparad för validering och ytterligare testning av modellerna, som ger möjlighet till vidareutveckling.

Populationernas utfallsfördelningar illustreras i Figur 2. De nyinskrivna har i betydligt större utsträckning kortare tid till arbete/studier – detta hänger ihop med det samband som syntes i Figur 1: i sökandestocken finns en blandning av olika

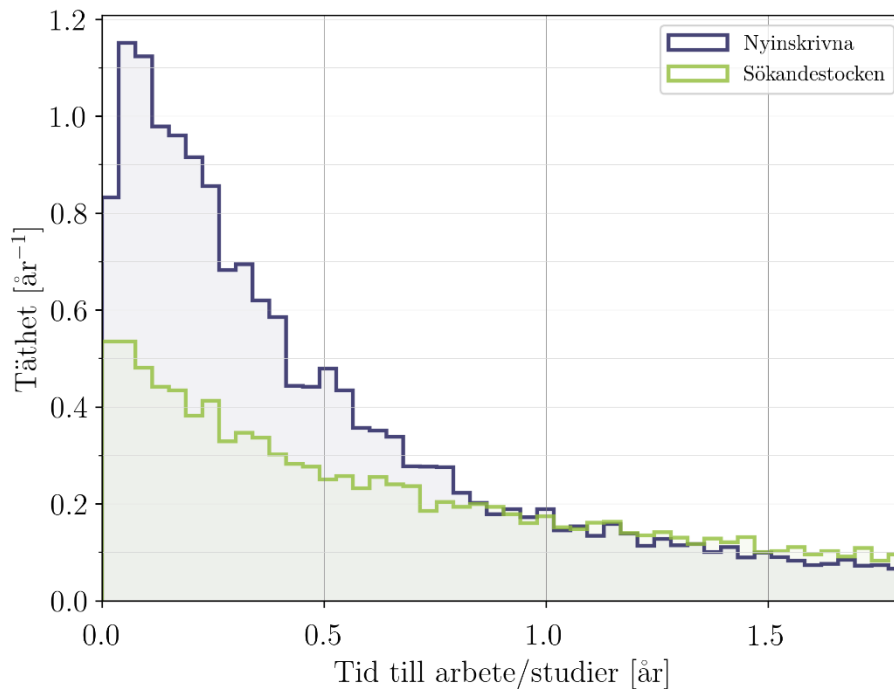
¹⁰ Motsvarande modeller för 2016 och 2017 studeras också i vissa delar av denna text, men om inget annat nämns är det 2018 som gäller.

¹¹ Man hade kunnat tänka sig att ta *ett* datum (säg, 1 mars 2018) och titta på alla inskrivna vid det datumet, men den här definitionen används för att undvika eventuella specificiteter för någon viss tid på året.

¹² Med *startdatum* menas det datum då bedömningen "görs"; den data som används för en viss individ i bedömningen och modellbyggandet gäller detta datum. Det är också utifrån detta datum som utfallet räknas – det gäller alltså *ytterligare* tid till avaktualisering till arbete eller studier (inte nödvändigtvis tid från inskrivning).

inskrivningstider, och längre inskrivningstid korrelerar med längre ytterligare arbetslöshetstid.

Figur 2. Histogram som visar fördelningarna för tid till arbete/studier för de olika träningspopulationerna för 2018 som används i denna undersökning.¹³



3.2 Avaktualiserade med okänd orsak och inskrivningar från före 2012 rensas bort

Omkring 20 procent av alla avaktualiseringar är av okänd orsak. Detta försämrar kvaliteten på utfallet eftersom den egentliga orsaken till avaktualisering kan vara av vitt skilda orsaker: antingen för att den sökande har fått ett arbete på egen hand eller för att hen har lämnat arbetskraften. För att inte detta ska komplicera undersökningen i denna rapport i onödan (genom att fördelningen och sammansättningen av dessa individer är olika i olika populationer, till exempel) rensas individer som har sådan avaktualisering före de avaktualiseras till arbete/studier bort ur hela undersökningen.¹⁴

Utöver dessa avaktualiseringar tas inskrivningar som påbörjades före augusti 2012 bort eftersom stora förändringar i yrkesdata gjordes då.

¹³ Den kortaste uppföljningstiden är 1,8 år; därför är tider som är längre än detta (i de flesta fall är de fortfarande inskrivna) utelämnade. En större andel av sökandestocken har uppföljningstider som är längre än så; därmed är alltså en större andel av sökandestocken utelämnad ur figuren.

¹⁴ För en riktig profilering bör man överväga möjligheten att använda imputerade värden för denna typ av bortfall som i Nilsson (2016). I projektplanen kopplad till denna promemoria föreslås också att studera hur mycket träffsäkerheten skulle påverkas av bättre utfallsdata (inkomstuppgifter).

I slutändan består tränings- och testpopulationerna av drygt 40 000 individer för sökandestocken och drygt 25 000 för de nyinskrivna (och för sökandestocken som nyinskrivna): de exakta värdena syns i Bilaga 3.

4 Träffsäkerheten definieras som en viktad andel som prediceras ”rätt”

4.1 Definition och motivering av träffsäkerhetsmålet

I denna undersökning jämförs prediktioner som är gjorda baserat på data som fanns tillgänglig vid ett visst tillfälle med de utfall som samma individer senare fått. Även när man har sådan möjlighet att jämföra prediktioner och utfall är det inte självklart hur man ska definiera träffsäkerheten hos en modell. I till exempel Bennmarker, Carling och Forslund (2007) och Böhlmark, Lundström & Ornstein (2021) sammanfattas träffsäkerheten med hur stor andel som prediceras till rätt sida om en viss gräns, alltså summan av andelen som bedöms vara ”nära” och som sedan visar sig vara det (får ett utfall på samma sida om gränsen) och andelen som bedöms vara ”långt ifrån” och som också visar sig vara det.

I denna rapport används i grunden samma idé för att definiera träffsäkerhet men med viss modifikation. Anledningen är att tolkningen av ett sådant mått blir komplicerad när man gör jämförelser av träffsäkerhet mellan *olika populationer* – och sådana jämförelser är helt centrala i denna undersökning. Därför modifieras måttet ovan genom att vikta de arbetssökande på så sätt att alla populationer jämförs som om de har samma utfallsfördelning. Viktningen beskrivs snart något mer ingående, men till att börja med beskrivs kort varför tolkningen av det grundläggande träffsäkerhetsmålet är komplicerad.

Sammanfattningsvis är problemet att det blir väldigt utslagsgivande hur utfallen är fördelade i förhållande till gränsen. För det första, ju större skillnaden är mellan antalet utfall på var sida om gränsen, desto lättare blir det att få en hög träffsäkerhet. Till exempel, om 90 procent av populationen står under gränsen fås 90 procent träffsäkerhet för en bedömning som säger att *alla* står nära arbetsmarknaden. En helt slumpmässig bedömning med 90 procent sannolikhet för ”nära”, skulle i samma situation få en träffsäkerhet på 82 procent. Båda dessa siffror är, trots deras naiva grund, betydligt högre än de som fås i Bennmarker, Carling och Forslund (2007) och Böhlmark, Lundström & Ornstein (2021), vilket belyser orimligheten i att jämföra dem.

En möjlig lösning av *detta* problem är att låta gränsen vara vid en viss kvantil i fördelningen av utfall för respektive population; exempelvis vid medianen. I fallet med medianen skulle en slumpmässig 50/50-bedömning ge en träffsäkerhet på 50 procent för alla populationer – så i viss mening skulle jämförelserna bli rättvisa. Med denna lösning kvarstår tyvärr ett problem: olika stor andel av fördelningen kan ligga

”nära” gränsen. Betrakta till exempel fördelningarna för sökandestocken och nyinskrivna i Figur 2: om gränsen går vid 35:e percentilen (medianen är censurerad för sökandestocken) för de olika fördelningarna ligger den vid ca 0,4 år för de nyinskrivna och vid ca 1,4 år för sökandestocken. För de nyinskrivna skulle denna gräns ligga relativt nära en stor andel av utfallen (med relativt stor risk för att hamna på fel sida gränsen), medan gränsen för sökandestocken skulle ligga på betydligt tryggare avstånd ifrån tyngre delar av fördelningen.

För att kompensera för detta i träffsäkerhetsmättet så viktas¹⁵ fördelningen för de olika populationerna om till samma fördelning när träffsäkerheten beräknas. Alltså, först byggs modellerna och sedan tas prediktioner fram. Efter detta viktas paren av prediktioner och verkliga utfall om till en referensfördelning baserat på utfallet. Slutligen justeras gränsen för de predicerade värdena så att en lika stor viktad andel av de predicerade värdena för träningspopulationen ligger under gränsen som för utfallen.¹⁶ Den viktade andelen som blivit predicerade på rätt sida om gränsen utgör det träffsäkerhetsmätt som används genomgående i denna text, med undantag för avsnitt 5.1.

Som referensfördelning används medelvärdet av fördelningarna för träningspopulationerna för sökandestocken och nyinskrivna, för att göra jämförelsen dem emellan så rättvis som möjligt. Denna referensfördelning beskrivs mer i Bilaga 2 och illustreras i Figur 3. Resultaten ändras inte dramatiskt för olika, någorlunda rimliga, referensfördelningar.¹⁷ Gränsen sätts till 365 dagar: ett naturligt värde som inte är alltför långt ifrån medianen för någon av fördelningarna. För referensfördelningen ligger 41,7 procent under denna gräns, vilket ger att slumpen får en träffsäkerhet på 51,4 procent med denna definition.

Träffsäkerhetsmättet i denna text kan ses som ett försök att jämföra under vilka förutsättningar en viss sammansättning av individer kan bedömas på ett mer respektive mindre träffsäkert sätt. Denna sammansättning av individer bestäms av deras utfall via referensfördelningen. Detta skiljer sig från att studera vilken sammansättning av individer som för någon viss gräns ger goda förutsättningar för hög träffsäkerhet. I en verklig situation så kan andra, och möjligen mindre ”rättvisa” träffsäkerhetsmätt vara mer relevanta. Det ”orättvisa” i att någon fördelning är mer utmanande kan i högsta grad vara viktigt: till exempel är det lättare att skilja ut de som står väldigt nära eller väldigt långt ifrån (fördelningens ”svansar”), så en sådan bedömning kan få en hög träffsäkerhet. Detta belyses också av träffsäkerheterna som redovisas i O’Connell, McGuinness & Kelly (2012) som studerar ett liknande verktyg i Irland. Där redovisas bland annat en träffsäkerhet på 84 procent: denna siffra gäller just en bedömning som ska skilja de som står allra närmast från de som står allra

¹⁵ Viktningen beskrivs i Bilaga 2.

¹⁶ Denna justering motsvarar en transformering av prediktionerna så att de *viktade fördelningarna* är lika för utfall och prediktioner för träningspopulationen. Utan en sådan justering skulle träningspopulationernas likhet med referenspopulationen påverka resultaten mer än nödvändigt.

¹⁷ Resultaten i avsnitt 5 har också studerats med referensfördelningar lika med var och en av träningspopulationerna. Träffsäkerheterna förändras lite (liksom slumpens), men förhållandena mellan de olika träffsäkerheterna blir i stort sett desamma.

längst ifrån¹⁸. Artikelförfattarna kommenterar att ”beslutsfattarna behöver göra avvägningen mellan täckningsgrad och träffsäkerhet”.

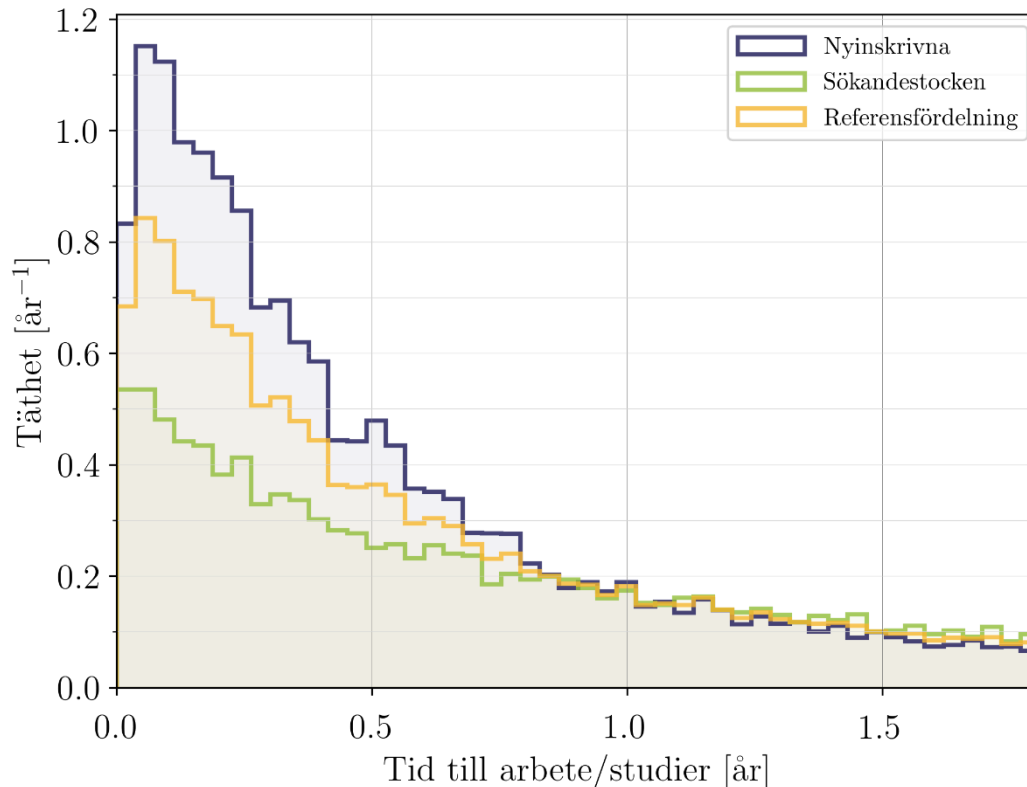
4.2 Hur träffsäkerheten beräknas – steg för steg

Sammanfattningsvis görs följande steg när träffsäkerheten beräknas:

1. En modell tas fram baserat på aktuell träningspopulation
2. Modellen används till att skatta avståndet från arbetsmarknaden för aktuell *testpopulation*
3. Testpopulationens utfall och skattningar ges vikter så att utfallen matchar referensfördelningen (se figur 3).
4. Andelen korrekta prediktioner beräknas för den viktade testpopulationen: alltså den viktade andelen arbetssökande för vilka utfall och prediktioner av arbetslöshetstiden ligger på samma sida om vald gräns. Denna viktade andel utgör den redovisade träffsäkerheten.

¹⁸ I fallet med 84 procents träffsäkerhet i O’Connell, McGuinness & Kelly (2012) skiljs de som bedöms ha mer än 80 procents sannolikhet till jobb inom ett år från de som bedöms ha mindre än 20 procents sannolikhet för detta. Endast ca 10 procent av deras studiepopulation befinner sig i var och en av dessa två svansar – de skiljer alltså ut de 10 procent närmast arbetsmarknaden från de 10 procent som står längst ifrån.

Figur 3. Fördelningarna för sökandestocken 2018 och nyinskrivna 2018 tillsammans med den referensfördelning som används i träffsäkerhetsmättet – som Figur 2 men inklusive referensfördelningen.



5 Resultat

5.1 Träffsäkerheten är relativt god vilket gör undersökningen relevant

I detta avsnitt (5.1) jämförs träffsäkerheten för modellerna i denna rapport med liknande träffsäkerheter i två andra studier. Eftersom vi inte kan använda någon referensfördelning (se avsnitt 4) i de andra studierna används inte det heller för våra resultat – just för *denna* jämförelse. Som diskuterats i avsnitt 4 är det inte helt oproblematiskt att göra sådana jämförelser, men värdena nedan indikerar ändå att modellen som används i denna undersökning har en rimligt god träffsäkerhet vilket bör göra att resultaten i denna rapport är relevanta.

I Böhlmark, Lundström & Ornstein (2021) studeras träffsäkerheten i Krom och där fås en träffsäkerhet på 68 procent. Med det utfall och den gräns som används där ger slumpen en träffsäkerhet på 57 procent. För att försöka jämföra med den träffsäkerheten kan gränsen i denna undersökning justeras så att slumpen ger samma

träffsäkerhet som där: den justeras då till 394 dagar. Träffsäkerheten blir för modellen tränad på sökandestocken blir då 73,9 procent - alltså högre än i Krom.

På liknande sätt kan träffsäkerheten för modellen tränad på nyinskrivna jämföras med den i Bennmarker, Carling och Forslund (2007), som också gäller nyinskrivna. Detta genom att flytta gränsen till medianen så att slumpen ger en träffsäkerhet på 50 procent. I Bennmarker, Carling och Forslund (2007) gäller detta ("ungefär" hälften får ett positivt utfall), och de får en träffsäkerhet på 68 procent. I denna undersökning fås, med medianen som gräns (301 dagar), 66,7 procent för de nyinskrivna – alltså aningen lägre än i Bennmarker, Carling och Forslund (2007).

5.2 Skillnaden i träffsäkerhet mellan sökandestocken och nyinskrivna är stor

Träffsäkerheten definierad enligt avsnitt 4 redovisas för de två populationerna i Tabell 1. Träffsäkerheten visas också för något olika varianter på modellen, där vissa variabler är exkluderade eller modifierade för att studera deras betydelse för träffsäkerheten. För de nyinskrivna saknas värden på vissa rader i tabellen eftersom dessa för nyinskrivna sammanfaller med närmast ovanstående rad som är ifylld.

Den första raden i tabellen är allra mest central; här jämförs de fullständiga modellerna med varandra. Det framgår att en träffsäkerhet på över 71 procent nås för sökandestocken medan de nyinskrivna får en träffsäkerhet på drygt 67 procent. Det skiljer drygt 4 procentenheter i träffsäkerhet mellan sökandestocken och den nyinskrivna. Det är svårt att generellt uttala sig om betydelsen av 4 procentenheters ökad träffsäkerhet men det kan åtminstone ställas i relation till skillnaden i träffsäkerhet mellan modellerna och slumpen: 20,1 procentenheter för sökandestocken. Drygt en femtedel av denna skillnad mot slumpen försvinner alltså för de nyinskrivna. I nästa avsnitt diskuteras fler av raderna i Tabell 1.

Tabell 1. Träffsäkerheten för de olika populationerna för olika varianter på modellen. Slumpen ger en träffsäkerhet på ca 51,4 procent.

	Sökandestocken	Nyinskrivna
1. Hela modellen	71,5 %	67,4 %
2. Utan variabeln inskrivningstid	70,7 %	
3. Utan historik på AF från tidigare inskrivningsperioder	71,4 %	66,5 %
4. Utan historik på AF	71,0 %	
4. Utan historik på AF och utan inskrivningstid	70,4 %	

6. Variabelvärden från inskrivningsdatumet 65,6 %

5.3 Information ökar och uppdateras under inskrivningen vilket kan förklara skillnaden mellan nyinskrivna och sökandestock

Raderna 2–5 i Tabell 1 visar resultat där vissa variabler – kopplade till tiden som inskriven – tagits bort ur modellen. Härifrån går det att utläsa att det går att ta bort flera sådana variabler ur modellen utan att träffsäkerheten påverkas särskilt mycket: om man tar bort alla ”historikvariabler”¹⁹ och variabeln inskrivningstid får man fortfarande en träffsäkerhet på 70,4 procent för sökandestocken – alltså bara en dryg procentenhet lägre än för modellen med dessa variabler. Detta kan jämföras med de drygt fyra procentenheterna i skillnad mot träffsäkerheten för nyinskrivna. Variabeln inskrivningstid i sig verkar vara ungefär lika viktig som historikvariablerna. Det kan vara värt att nämna att inskrivningstiden är en av de absolut mest signifikanta variablerna i modellen när den ingår – ändå kan den till stor del ersättas av andra variabler²⁰.

Avsaknaden av information om inskrivningstiden och ”historikvariablerna” förklarar alltså långt ifrån hela minskningen i träffsäkerhet när man går från sökandestock till nyinskrivna. Däremot indikerar rad 6 i Tabell 1 att skillnaden i träffsäkerhet beror på att variablerna generellt sett blir av högre kvalitet eller mer informativa under tiden som inskriven. På rad 6 har nämligen variabelvärdena från inskrivningsdatumet använts, och då går träffsäkerheten ned till under 66 procent – alltså en bit under träffsäkerheten för de nyinskrivna! En del av denna skillnad består i inskrivningstid och historikvariabler, men det är också många andra variabler som förändras under tiden som inskriven. I avsnitt 5.6 framgår att uppdaterade värden på sökandekategorier är allra viktigast.

Att träffsäkerheten påverkas så här mycket av tillkommande information under inskrivningstiden kan tyda på att det finns stor potential till ytterligare ökad träffsäkerhet genom tillgång till än mer information. För att inte överskatta denna potential kan det vara värt att notera att den information som visar sig viktig i denna rapport är av en särskild natur: den tillkommer under inskrivningstiden och bär därmed med sig information om händelser och beslut som rör den arbetssökande under tiden som inskriven – alltså på väg mot ett eventuellt utfall. Sådan information kan eventuellt vara betydligt mer värd än mer detaljer om den arbetssökande vid

¹⁹ ”Historikvariabler” syftar på variabler som beror av *tidigare* värden på uppgifter som är registrerade om den arbetssökandes tid som inskriven (även under tidigare inskrivningar). Mer detaljer återfinns i bilaga 2.

²⁰ Om man tillhör etableringsmålgruppen eller ej, kön och funktionshinderkod har stor betydelse i alla modeller och betydelsen av dessa viktiga variabler förstärks i allmänhet när inskrivningstiden tas bort. Inskrivningstiden har störst betydelse för svenskfödda (se bilaga 1), där den interna träffsäkerheten går ned klart mer än för helheten när den tas bort (närmare en procentenhet). Flera ytterligare variabler ökar tydligt i betydelse för denna grupp när inskrivningstiden tas bort är till exempel olika sökandekategorier, främst sådana som gäller deltid eller tillfälliga anställningar.

inskrivningen.²¹ Däremot har bättre utfallsdata i form av inkomstuppgifter en särställning eftersom den stora mängden avaktualiseringar av okänd orsak ger ett betydande bortfall som inte är slumpmässigt fördelat – det är inte ytterligare en förklarande variabel i mängden.

5.4 Träffsäkerheten påverkas mer av testpopulationen än av träningspopulationen

Detta avsnitt undersöker hur robusta modellerna är för att användas för en annan population än vilken den tränats för. Dels testas till exempel modellen tränad för nyinskrivna på sökandestocken och vice versa, dels används också tränings- och testpopulationer från olika år: 2016, 2017 och 2018. Dessutom inkluderas den variant på sökandestocken 2018 som har variabelvärden från inskrivningen (rad 6 i Tabell 1), både som träningspopulation och testpopulation.

Träffsäkerheter för alla kombinationer av tränings- och testpopulationer syns i Tabell 2. Notera att resultaten längs diagonalen för 2018 är samma som de på översta raden och i nedre vänstra hörnet i Tabell 1: längs diagonalen används nämligen tränings- och testpopulationer som är slumpmässigt dragna från samma population, precis som i Tabell 1.

Flera intressanta mönster framträder i Tabell 2. Till att börja med: modeller tränade på data från ett av dessa år fungerar i det närmaste lika bra på testpopulationer från andra år. Till exempel ger modellen tränad på sökandestocken 2016 en träffsäkerhet på 71,6 procent för sökandestocken 2018, alltså en aning högre än modellen som också är tränad på sökandestocken 2018. Liknande mönster syns för andra kombinationer av år, och även för nyinskrivna. Detta är viktigt och intressant eftersom detta ligger nära en verklig användning²². Notera att detta inte utan vidare går att generalisera till andra år; om det varit större skillnader i konjunkturläge mellan åren skulle resultaten kunna se väldigt annorlunda ut.

Ett annat väldigt intressant mönster är att testpopulationen är viktigare än träningspopulationen: till exempel fås en träffsäkerhet på mellan 69,0 procent och 71,7 procent för sökandestocken 2018 oavsett vilken av träningspopulationerna som

²¹ Inte minst antyder den till synes lägre träffsäkerheten för verktyget i KROM, som använder mer detaljerad individdata, att detta inte behöver vara det allra viktigaste. Ett annat exempel på att mer detaljerad individdata kan ha begränsad betydelse återfinns i Ornstein och Thunström (2021): där undersöks bland annat hur tillgången till handläggares bedömningar av vissa egenskaper påverkar träffsäkerheten, och givet grundläggande information ökar inte detta träffsäkerheten nämnvärt. Dessutom har verktyget i Ornstein och Thunström (2021) liknande träffsäkerhet som verktyget i KROM, trots att det är betydligt enklare. Ytterligare ett exempel som belyser att tillgång till en stor mängd bakgrundsdata inte alltid är avgörande återfinns i Salganik m fl. (2020). Där får en stor mängd team uppgiften att förutsäga olika viktiga händelser några år framåt i olika familjers livsbanor. Teamen har tillgång till en omfattande mängd bakgrundsdata som rör familjerna och de använder mer eller mindre avancerade AI-modeller för att förutsäga hur det ska gå. Dels visar det sig att det är väldigt svårt att lösa uppgiften, men det visar sig också att betydligt enklare modeller med ett mycket litet antal variabler utvalda av experter är nästan lika träffsäkra. För övrigt har denna studie även bäring på avsnitt 5.4 – att träffsäkerheten påverkas mer av testpopulationen än av träningspopulationen: precis som i avsnitt 5.4 får de olika modellerna i Salganik m fl. (2020) resultat som är väldigt lika varandra och de har därmed typiskt sett fel på samma familjer.

²² För att vara av praktisk nytta behöver profileringarna göras utan tillgång till verkliga utfall samtidigt som modellbygget kräver verkliga utfall, vilket kräver en viss uppföljningstid. Alltså krävs en viss tidsförskjutning mellan träningspopulation och den population som profileras i praktiken.

använts, medan modellen som använder sökandestocken 2018 som *träningpopulation* når träffsäkerheter mellan 65,3 procent och 72,2 procent för de olika testpopulationerna. Faktum är att denna modell – alltså tränad på sökandestocken 2018 – når högre träffsäkerhet på testpopulationer för sökandestocken från både 2016 och 2017 än för sökandestocken 2018. I linje med detta mönster framgår det att till och med modeller som är tränade på nyinskrivna får högre träffsäkerhet på testpopulationer ur sökandestocken än på testpopulationer av nyinskrivna; de olika modellerna byggda på nyinskrivna får nästan lika hög träffsäkerhet för sökandestocken 2018 som modellen för sökandestocken 2018 utan historikvariablerna och inskrivningstid fick i Tabell 1. Mönstret gäller även för sökandestocken 2018 med variabelvärden från inskrivningen: den ouppdaterade informationen ger betydligt sämre träffsäkerhet när det gäller testpopulationen medan träffsäkerheten påverkas mindre av att *tränas* på uppdaterad information (och sedan användas på uppdaterad information).

Tittar man närmare på prediktionerna för en viss testpopulation framgår att de olika träningspopulationerna ger liknande avståndsbedömningar: för en given testpopulation fås samma resultat för 83–92 procent av prediktionerna²³ med de andra träningspopulationerna (jämfört med träningspopulationen som motsvarar testpopulationen). Med detta i åtanke är det inte förvånande att det är testpopulationen som avgör.

²³ Med samma resultat menas här att prediktionerna ligger på samma sida om gränserna som används i träffsäkerhetsmättet. Andelarna 83-92 procent är viktade enligt referensfördelningen som används i träffsäkerhetsmättet; skillnaden mot oviktade andelar är bara någon procentenhet.

Tabell 2.. Raderna visar vilken population som använts för att bygga modellen och kolumnerna visar vilken testpopulation som avses.

		Testpopulation						
		Sökande- stocken 2016	Nyinskriv- na 2016	Sökande- stocken 2017	Nyinskriv- na 2017	Sökande- stocken 2018	Nyinskriv- na 2018	Sökande- stocken 2018 med variabler fr. inskrivn.
Trän- ings- popu- lation	Sökande- stocken 2016	73,2%	68,7%	73,5%	68,3%	71,6%	65,8%	66,2%
	Nyinskrivna 2016	72,0%	71,2%	72,0%	69,7%	70,0%	67,2%	66,7%
	Sökande- stocken 2017	72,7%	67,7%	73,2%	68,0%	71,7%	65,6%	65,8%
	Nyinskrivna 2017	71,9%	70,7%	71,8%	69,9%	70,2%	67,3%	66,2%
	Sökande- stocken 2018	71,4%	65,9%	72,2%	67,1%	71,5%	65,3%	65,3%
	Nyinskrivna 2018	71,1%	69,1%	71,7%	69,2%	69,8%	67,4%	66,1%
	Sökande- stocken 2018 med variabler fr. inskrivn.	69,0%	65,9%	69,5%	66,7%	69,0%	65,2%	65,6%

5.5 Träffsäkerheten är inte allt

Träffsäkerheten är inte det enda som avgör kvaliteten hos en statistisk bedömning och ett viktigt exempel är systematiska fel, som inte behandlas i denna rapport²⁴. Den minskning i träffsäkerhet som trots allt uppstår genom att till exempel använda nyinskrivna som träningspopulation när sökandestocken är testpopulation är inte slumpmässigt fördelad: det är alltså ett *systematiskt* fel som påverkar de som har längre inskrivningstid på ett annat sätt än nyinskrivna. Trots att träffsäkerheten inte

²⁴ Det studeras dock, för Krom, i Böhlmark, Lundström och Ornstein (2021).

minskar särskilt mycket kan detta ändå leda till att olika grupper kan behandlas systematiskt olika på ett sätt som inte är önskvärt.

5.6 Uppdaterade sökandekategorier viktigaste informationstillskottet – följt av inskrivningstid

I avsnitt 5.3 konstateras att de uppdateringar av variabelvärden som görs under inskrivningstidens gång är avgörande för skillnaden i träffsäkerhet mellan sökandestocken och nyinskrivna. Det är då intressant att se *vilka variabler* som driver detta. För att undersöka detta sätts olika grupper av variabler hos sökandestockens testpopulation till sina respektive värden vid inskrivningsdatumet, grupp för grupp.²⁵

Resultaten, i form av förändring i träffsäkerhet, syns i Tabell 3. I tabellen syns också hur stor andel av individerna som fått förändringar i de respektive variabelgrupperna.

Inskrivningstiden och historikvariablerna, vars påverkan också studerades i avsnitt 5.3, *har* helt klart en viss påverkan, vilket är förväntat²⁶. Skillnaderna är något större än när dessa variabler utelämnades helt ur modellen; detta är naturligt eftersom modellen då kunde hitta andra variabler som kompenserade. Den enskilt viktigaste variabeln att ha uppdaterade värden för är ändå sökandekategorin: används inskrivningsdatumets värden för denna så går träffsäkerheten ned med 1,3 procentenheter.

Ej uppdaterade funktionshinderkoder slår i sig inte lika hårt på *helheten*, men denna skillnad kan vara nog så viktig för dem det berör: alla har en sökandekategori men bara vissa är aktuella för funktionshinderkod.²⁷ Uppdateringarna av variabler kopplade till sökta yrken och erfarenhet har begränsad betydelse, trots att de uppdateras i nästan lika stor utsträckning som sökandekategorierna.²⁸

²⁵ Eftersom det framgår i avsnitt 5.4 att testpopulationen är betydligt viktigare för träffsäkerheten än träningspopulationen så görs en förenkling i denna del av undersökningen: endast testpopulationens variabelvärden varierar och uppdaterade variabelvärden används genomgående för träningspopulationen.

²⁶ Mycket tyder på att arbetslöshetstid i sig ökar risken för framtida arbetslöshet (Kreuger m fl 2014, Eriksson & Lagerström 2004), vilket kan förklaras av två huvudanledningar: att arbetslösa över tid förlorar i humankapital (t ex tappat motivation och anknytning till arbetsmarknaden) och att perioder av arbetslöshet skickar negativa signaler till potentiella arbetsgivare. I vilket fall förklaras framtida arbetslöshet av arbetslöshetstid när man kontrollerar för många andra egenskaper (Kreuger m fl 2014).

²⁷ Endast 5 procent av sökandestocken hade funktionshinderkod vid inskrivningsdatumet medan 19 procent har det med de uppdaterade variablerna.

²⁸ Vissa variabler kopplade till sökta yrken och erfarenhet är bland de viktigare variablerna i modellen, men uppdateringar av dem är alltså inte av så stor betydelse. Det är däremot inte otänkbart att *huruvida* variablerna har uppdaterats eller ej skulle kunna användas som en bra förklarande variabel; justeringarna må vara små men *att de görs* kan säga en del om motivation och kännedom om arbetsmarknaden hos den sökande och/eller ett pågående matchningsarbete.

Tabell 3. Förändringen i träffsäkerhet för modell tränad på sökandestocken 2018 när en del av variablerna i testpopulationen sätts till sina värden vid inskrivningen,²⁹ samt hur stor andel av populationen som fått förändringar bland dessa variabler.

Variabler som vid inskrivning i testpopulation	Funktionshinderkod	Sökandekategori	Sökta yrken och erfarenhet	Historikvariabler	Inskrivningstid	Övriga
Skillnad i träffsäkerhet (procentenheter)	-0,25	-1,28	-0,27	-0,91	-0,95	-0,15
Andel med förändring	14%	64%	53%	88%	100%	100%

6 Sammanfattning av undersökningen och dess resultat

Träffsäkerheten hos en bedömning av arbetssökandes stödbehov reflekterar i hur stor utsträckning bedömningarna är korrekta – god träffsäkerhet är alltså en förutsättning för att kunna tilldela resurser på ett effektivt sätt. Historiskt sett har insatser främst riktats till arbetssökande med en längre pågående arbetslöshetstid, eftersom dessa i allmänhet har större risk för fortsatt arbetslöshet. Under senare år har insatser i ett tidigare skede efterfrågats i större utsträckning, och en statistisk bedömning av risken för längre arbetslöshet kan användas för att bedöma stödbehovet i ett sådant tidigt skede, liksom senare under arbetslöshetstiden.

I denna undersökning studeras i vilken utsträckning träffsäkerheten hos en statistisk bedömning av avståndet till arbetsmarknaden begränsas av att göra en statistisk bedömning i ett tidigt skede. Närmare bestämt jämförs träffsäkerheten som fås för olika populationer: sökandestocken (som har varierande inskrivningstid) och nyinskrivna. Därutöver undersöks hur träffsäkerheten påverkas av att modellen utvecklas på en population och används på en annan.

Jämförelser av träffsäkerheter för olika populationer blir lätt orättvisa och eftersom sådana jämförelser är centrala i denna undersökning uppstod ett behov av metodutveckling: avsnitt 6.1 sammanfattar det nya träffsäkerhetsmått som utvecklats och använts i undersökningen. I återstoden av detta avsnitt, avsnitt 6.2–6.6, sammanfattas undersökningens viktigaste resultat.

²⁹ Testpopulationen är i grunden sökandestocken 2018. Om *alla* variabler sätts till sina värden vid inskrivningen minskar träffsäkerheten med 6,2 procentenheter (klart mer än summan av skillnaderna i tabellen), vilket ger träffsäkerheten 65,3 procent som syns i Tabell 2 (femte raden, sista kolumnen).

6.1 Ett nytt träffsäkerhetsmått tas fram för att kunna jämföra träffsäkerhet mellan olika populationer

Jämförelser av träffsäkerhet för olika populationer är svåra att göra på ett rättvist sätt; därför definieras och används ett speciellt träffsäkerhetsmått. Måttet baseras på hur stor andel som predicerats korrekt, men är utformat så att olika populationer jämförs som om de hade samma sammansättning på utfallen. Detta är ett försök att isolera frågan om under vilka förutsättningar en viss sammansättning av individer kan bedömas på ett mer respektive mindre träffsäkert sätt. Jämförelserna bortser därmed med avsikt från hur lätt en viss fråga är att besvara för en viss population: till exempel att det kan vara lättare att urskilja långtidsarbetslösa i en population där detta är ovanligt.

6.2 Högst träffsäkerhet fås för sökandestocken

Genom att jämföra träffsäkerheten mellan sökandestocken och nyinskrivna söks den sammantagna betydelsen av varierande tid som inskriven för träffsäkerheten. Det visar sig vara stor skillnad i träffsäkerhet mellan dessa grupper: drygt fyra procentenheter. Detta motsvarar en dryg femtedel av skillnaden mot den träffsäkerhet en helt slumpmässig avståndsbedömningskulle ge – en betydligt högre träffsäkerhet nås alltså för sökandestocken än för nyinskrivna.

6.3 Sökandestockens högre träffsäkerhet beror till stor del på tillkommande information under tiden som inskriven

Om variabelvärdena från inskrivningsdatumet används för sökandestocken minskar träffsäkerheten så att den blir betydligt lägre – till och med lägre än träffsäkerheten för nyinskrivna. Detta pekar tydligt mot att skillnaden i träffsäkerhet mellan sökandestock och nyinskrivna i huvudsak kan förklaras av den ytterligare information som tillkommer under tiden som inskriven: bland annat är det viktigt att sökandekategorier och funktionshinderkoder uppdateras. Vissa av dessa förändringar kan vara rena rättelser, men de kan också innehålla information om händelser i eller omkring den arbetssökande under inskrivningstiden: vägval och bedömningar som gjorts av både den arbetssökande och arbetsförmedlare, till exempel.

Som väntat är själva inskrivningstiden också en av de variabler som är viktigast att uppdatera: i linje med andra studier finns ett tydligt samband mellan inskrivningstid och arbetslöshetstid, även när man kontrollerar för en stor mängd observerbara variabler. Inskrivningstiden är dock inte ensamt avgörande.

6.4 Skillnader i träffsäkerhet beror främst på vilken population modellen tillämpas på, oavsett vilken population modellen tränats på

För att förstå resultaten är det viktigt att skilja på den population som används för att bygga modellen (träningpopulationen) och den population som används när

träffsäkerheten mäts (testpopulationen). Genom att variera både tränings- och testpopulationer, ser man att det som framför allt har betydelse för träffsäkerheten är *testpopulationen*. De olika modellerna ger i stor utsträckning liknande prediktioner och har till stor del fel för samma individer.

Detta ligger också i linje med andra studier där människors framtid prediceras: det är snarare vissa individer som är svårbedömda och inte olika modeller som presterar särskilt olika. Att modellens utformning har en begränsad förmåga att öka träffsäkerheten avser både inverkan av ökad detaljnivå i den data som modellen bygger på och inverkan av att använda en mer avancerad modell.

6.5 Modell för nyinskrivna har hyfsad träffsäkerhet på sökandestocken

Det visar sig att en modell tränad på nyinskrivna inte har särskilt mycket lägre träffsäkerhet när den används på sökandestocken (jämfört med en modell tränad på sökandestocken); skillnaden är knappt två procentenheter. Det verkar kort sagt som att även den mer begränsade information som finns vid inskrivningstillfället räcker långt för att träna en modell.

6.6 Modell tränad på ett visst år går bra att använda på ett annat år

Utöver att modeller tränade på nyinskrivna testats på sökandestocken, har modeller tränade på data från ett år testats på data från ett annat. För de år som studeras, 2016–2018, observeras väldigt små skillnader i träffsäkerhet när träningspopulationen har ändrats. Detta är viktigt eftersom man i praktiken alltid behöver träna sin modell på data som är något äldre än den man vill använda modellen på: annars har man inte tillgång till kända utfall. Resultaten gäller ett litet antal näraliggande år utan alltför dramatiska förändringar på arbetsmarknaden; om större förändringar sker (som under en pandemi) kan bilden bli en annan.

7 Slutsatser och rekommendationer

Resultaten som sammanfattats i avsnitt 6 leder tillsammans med diskussionen i rapporten till följande slutsatser och rekommendationer.

7.1 Tillkommande information ger högre träffsäkerhet – överväg att avvakta med dyrare insatser och säkerställ datainsamling under inskrivningstiden

En statistisk modell har större möjlighet till hög träffsäkerhet för sökandestocken än för nyinskrivna, till stor del på grund av information som tillkommer under den tid

som en arbetssökande är inskriven på Arbetsförmedlingen: främst uppdaterad information om till exempel sökandekategorier men även själva inskrivningstiden. Detta kan ha implikationer på vilka typer av beslut som är lämpliga att ta för nyinskrivna respektive individer med längre inskrivningstid; kanske bör man, i likhet med traditionell arbetsmarknadspolitik, avvakta med vissa insatser (framför allt dyrare insatser med större risk för inlåsnings effekter) tills en bedömning gjorts vid en längre inskrivningstid – *hur lång tid* man ska vänta blir en komplicerad avvägning mellan träffsäkerheten och de negativa konsekvenserna av väntetid respektive inlåsnings och kostnader för onödiga insatser. De som står allra längst från arbetsmarknaden kan bedömas tillräckligt väl tidigare än de som står närmare mitten; möjligen redan vid inskrivningen. Avvägningen är i slutändan upp till beslutsfattarna, men bör också studeras närmare för att ge ett bättre underlag. Det är också viktigt att myndigheten säkerställer att information även framöver inkommer under tiden som inskriven.

7.2 Undersök gärna möjligheten till ökad träffsäkerhet, men överskatta den inte

Att det uppstår en stor skillnad mellan sökandestock och nyinskrivna på grund av att mer information finns tillgänglig för sökandestocken kan ses som en indikation på att än mer tillkommen information kan förbättra träffsäkerheten ytterligare. Denna potential kanske inte ska överskattas eftersom den viktiga informationen som tillkommer är av en speciell karaktär: den gäller händelser och beslut under den sökandes tid som inskriven. Likväl kan det vara av stort värde att undersöka värdet av att samla in ytterligare uppgifter, särskilt då under pågående arbetslöshetstid. De uppgifter som framkommer som viktigast är sådana som är beroende av dokumenterade bedömningar, utredningar och beslut inom ramen för Arbetsförmedlingens verksamhet. Mer detaljerad allmän information om de arbetssökande vid inskrivningen, liksom mer avancerad modellering, har sannolikt inte lika stor betydelse.

Tillgång till bättre utfallsdata – förhoppningsvis i form av inkomstuppgifter – har en särställning. Detta eftersom avaktualiseringar med okänd orsak i nuläget innebär ett ej slumpmässigt *bortfall*.

7.3 Studera hur möjligheten till träffsäkerhet beror på frågan som ställs

För att åstadkomma en rättvis jämförelse av träffsäkerheten för olika populationer används här ett träffsäkerhetsmått som minskar betydelsen av den fråga som ställs: till exempel var gränsen går för vilka som anses behöva stöd i förhållande till utfallsfördelningen för populationen det gäller. Det går också att tänka sig något annorlunda frågeställningar, som att man vill skilja ut dem som står allra längst ifrån arbetsmarknaden från dem som står allra närmast. Vilken fråga som ställs blir väldigt viktigt för hur bra träffsäkerhet man kan nå i den praktiska användningen av en

statistisk bedömning. Detta beroende – som bara berörs flyktigt i denna rapport – borde undersökas och beskrivas mer ingående.

7.4 Träffsäkerhet är inte allt – studera även systematiska fel när modeller används för en annan population

En modell byggd på nyinskrivna kan få hög träffsäkerhet även när den används på sökandestocken – nästan lika hög som en modell byggd på sökandestocken. Detta är ett exempel på att träffsäkerheten beror mer på vilken population som modellen används på än vilken population som använts i modellbygget. Man bör ändå vara försiktig med att använda en modell som är utvecklad för en annan typ av population även om träffsäkerheten kan bli hög; i detta fall behandlas arbetsökande med olika inskrivningstid systematiskt olika vilket kan göra att olika grupper av arbetsökande behandlas olika på ett oönskat sätt. Exempelvis kommer en modell baserad på nyinskrivna sannolikt systematiskt överskatta jobbchansen hos arbetsökande med lång inskrivningstid. Kraftiga konjunktursvängningar kan också medföra liknande konsekvenser. Systematiska fel – till skillnad från träffsäkerhet på helheten – som konsekvens av byte av population har inte beaktats i denna rapport, men bör undersökas och beskrivas.

7.5 Modell från studerade år går att använda på andra år – studera detta under mer dramatiska förändringar

På liknande sätt kan en modell byggd på data från ett visst år användas under andra år utan att förlora mycket i träffsäkerhet, vilket är mycket viktigt för praktisk användning. Detta gäller åtminstone för 2016–2018, någorlunda näraliggande år med liknande konjunkturläge. Ytterligare en rekommendation är att studera sådana byten av populationer under mer dramatiska förändringar. Även en sådan undersökning bör beakta systematiska fel och dess konsekvenser.

Referenser

Arbetsförmedlingen (2020a). *Beskrivning av arbetsmarknadspolitisk bedömning med ett statistiskt bedömningsstöd i Kundval Rusta och Matcha 1.0*. Af-2019/0038 1388.

Arbetsförmedlingen (2020b). *Arbetsförmedlingens verksamhetsplan 2020, reviderad 2020-06-12*. Af-2019/0017 2992.

Arbetsmarknadsdepartementet (2013). *Uppdrag till Arbetsförmedlingen avseende genomförande av tidiga insatser*. Regeringsbeslut A2013/2678/A.

Benmarker H., Carling K. & Forslund A. (2007). *Vem blir långtidsarbetslös?* (IFAU Rapport 2007:20).

Böhlmark A., Lundström T. & Ornstein P. (2021). *Träffsäkerhet och likabehandling vid automatiserade anvisningar till insatser inom KROM*. Arbetsförmedlingen, PM under bearbetning.

Eriksson S. & Lagerström J. (2007). *Väljer företag bort arbetslösa arbetssökande?* (IFAU Rapport 2004:3).

Hartman L. (2018). *Profileringsverktyg och tillitsbaserad styrning – synergi snarare än motsättning*. I *Framtidens arbetsförmedling*, (red.) L. Calmfors och A. Bergström, Stockholm: Fores, s. 201-237.

Kreuger A.B., Cramer J. & Cho D. (2014). *Are the Long-Term Unemployed on the Margins of the Labor Market?* *Brookings Papers on Economic Activity*, Spring 2014, pp. 229-299.

Nilsson P. (2016). *An Imputation Model for Dropouts in Unemployment Data*. *The Journal of Official Statistics*, 32(3), 719–732.

O’Connell P.J., McGuinness S. & Kelly E. (2012). *The Transition from Short- to Long-Term Unemployment: A Statistical Profiling Model for Ireland*. *The Economic and Social Review*, 43(1), 135–164.

Ornstein P. & Thunström H. (2021). *Träffsäkerhet i bedömningen av arbetssökande. En jämförelse av arbetsförmedlare och en statistisk modell*. Arbetsförmedlingen, rapport under bearbetning.

Salganik et al. (2020). *Measuring the predictability of life outcomes with a scientific mass collaboration*. *Proceedings of the National Academy of Sciences*, 117(15), 8398–8403

Bilaga 1. Lite mer om modellen

6 undermodeller används, med alla kombinationer av: kön, varit i etableringsmålgrupp/annat uppehållstillstånd/inget uppehållstillstånd

Eftersom förutsättningarna skiljer sig ganska mycket för olika grupper av arbetssökande görs separata regressioner för sex utvalda grupper. De arbetssökande delas in i tre grupper beroende på om de är svenskfödda, annars om de tillhör eller har tillhört etableringsmålgruppen (främst flyktinginvandrare och deras anhöriga) eller om de har ett annat skäl till uppehållstillstånd (främst arbetskraftsinvandring eller invandrade som anhöriga till svenskfödda). Var och en av dessa grupper delas sedan in efter kön. För varje grupp görs en regression och efterföljande transformering enligt beskrivningen i 6.2-6.4.

Individkaraktäristika och historik på Arbetsförmedlingen används som förklarande variabler

För varje undermodell sker en linjär regression enligt 6.3, utan interaktioner, och med förklarande variabler enligt tabell Tabell 4 i de fall de är relevanta³⁰. För en del variabler, t ex för födelseland eller sökandekategorier, används en dummyvariabel för alla värden som förekommer minst 100 gånger – antalet resulterande dummyvariabler för modellen för kvinnor utan uppehållstillståndsdata (främst svenskfödda) i sökandestocken 2018 syns i tabellen (" + 1" innebär ytterligare en implicit variabel: t ex för utbildningsnivå förekommer dummies för nivå 1-6, om utbildningsnivån är 0 så är ingen av dessa dummies lika med 1). Innan regressionen tas variabler bort så att variablerna blir linjärt oberoende för de ocensurerade observationerna (så få variabler som möjligt tas bort). De variabler som är märkta med en asterisk kallas i denna rapport för historikvariabler: de är sådana som använder tidigare värden på variabler i Arbetsförmedlingens register. De flesta av dessa variabler kan gälla både tidigare och nuvarande inskrivning, och modeller som inte har med historik från tidigare inskrivningar studeras också. Notera att kön och uppehållstillståndstyp ingår genom att de ingår i definitionen av undermodellerna.

Tabell 4. Asterisk utmärker "historikvariabler": de använder tidigare värden på variabler Arbetsförmedlingens register.

Dummyvariabler	Antal (för vissa dummyvariabler)
Födelseland	12
Kommun	26
Län	20
Sökandekategori hos AF	13
Sökta yrken med erfarenhet (SSYK)	22
Antal sökta yrken	4 + 1
Antal sökta yrken med erfarenhet	4 + 1
Utbildningsnivå	6 + 1
Har önskat arbete	
Haft arbete via AF*	

³⁰ Till exempel ingår inte tid med uppehållstillstånd i undermodellen för svenskfödda, och inskrivningstid ingår inte i modellerna för nyinskrivna.

Avbrutit arbete via AF *

Haft AUB*

Avbrutit AUB *

Haft FUB*

Avbrutit FUB*

Avbrutit arbete via AF*

Avbrutit praktik via AF*

Har funktionshinderkod

Antal tidigare inskrivningsepisoder* 14 + 1

Senaste avaktualiseringsorsak* 6 + 1

Andragradspolynom

Ålder

Tid m uppehållstillstånd

Inskrivningstid

Tiden till arbete eller avaktualisering till studier räknas

Tiden till den första avaktualiseringen till arbete eller studier används som utfallsmått; i fortsättningen kallas detta ”tid till arbete/studier”. *Logaritmen* på denna tid används i regressionerna, på grund av erfarenhet av liknande utfallsvariabler, fördelningen på utfallen, och eftersom tiden till arbete eller studier måste vara större än noll³¹.

Ett alternativ till att använda tid till arbete/studier är att använda motsvarande status (avaktualiserad eller ej) vid en *viss tid* sedan startdatumet, men då förlorar man information. Ytterligare ett alternativ, som används i Bennmarker, Carling och Forslund (2007), är att använda Hazardmodeller; detta görs ej här eftersom det

³¹ Ett villkor för att vara med i stickprovet är att man inte är i arbete vid startdatumet – således kan man inte vara i arbete förrän tidigast dagen efter.

kräver en mer komplicerad implementering och eftersom vi vill hålla modellen enkel för överskådlighet.

Eftersom tiden som har förflutit mellan startdatumen och datainsamlingen är begränsad³² så är det många arbetssökande som ännu inte har avaktualiserats till arbete eller studier; vi säger att utfallen i dessa fall är *censurerade*. För dessa personer vet vi inte hur många dagar de har till arbete/studier – däremot vet vi att det har tagit *minst* så lång tid som uppföljningstiden är för denna person. Censureringen hanteras med en Tobit-modell³³, vilket gör att vetenskapen om att det minst har tagit så lång tid verkligen används.

Transformering av fördelningen

Fördelningen för skattningarna som resulterar från regressionerna ovan blir inte fördelade precis som utfallen för träningspopulationen: skattningarna tenderar att bli färre i ”svansarna”, alltså för de längsta och kortaste arbetslöshetstiderna. En transformering av resultaten görs därför för varje delmodell så att *fördelningen* när utfallen för träningspopulationen skattas blir samma som fördelningen för själva utfallen (där inte censureringen är aktuell). Den inbördes ordningen för skattningarna i varje delmodell påverkas inte.

För träningspopulationen kan detta åstadkommas genom att ordna alla utfall i stigande ordning och sedan transformera skattningarna genom att tilldela dem dessa värden utefter skattningarnas ordning: den lägsta skattningen tilldelas värdet för det lägsta utfallet, och så vidare. För att göra detta mer generellt görs det genom att beräkna lika många kvantiler för träningspopulationen som det finns observationer i testpopulationen: kvantilerna för $0, 1/(n_{test}-1), 2/(n_{test}-1), \dots, (n_{test}-2)/(n_{test}-1), 1$. Sedan tilldelas dessa kvantiler till skattningarna i enlighet med skattningarnas ordning, alltså så att ordningen på skattningarna bibehålls efter transformeringen. På detta sätt får skattningarna för testpopulationen en fördelning som är mycket lik den för träningspopulationens utfall.

Arbetslöshetstidens beroende av inskrivningstid

I huvudtexten diskuterades framför allt inskrivningstidens betydelse för *träffsäkerheten*. Det är också intressant att kort studera inskrivningstidens betydelse för själva avståndet till arbetsmarknaden i lite mer detalj. Redan i Figur 1 syns att arbetssökande med längre inskrivningstid i genomsnitt har längre ytterligare tid till arbete/studier. I den figuren syns detta endast som det samband det är – inga försök att skilja inskrivningstiden från dess samvariation med andra egenskaper hos de arbetssökande har gjorts där. I Figur 4 syns den skillnad i predicerad tid till

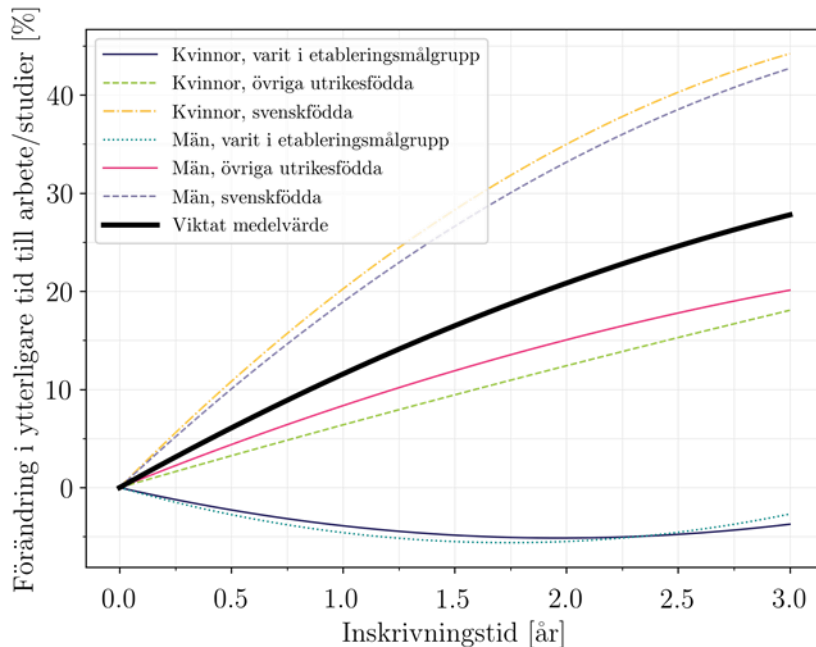
³² Utfallsvariablerna har uppdaterats 8 september 2020 så att den kortaste uppföljningstiden är 618 dagar (för individer med startdatum 31 dec 2018).

³³ Tobitmodellen hanterar censureringen genom att modifiera den ”vanliga” likelihood-funktionen så att den inte bara består av täthetsfunktionen för observationerna (givet parametrarna), utan även sannolikheten för observationer att ligga i det censurerade området (givet parametrarna). Tobitmodellen antar homoscedastiska normalfördelade fel.

arbete/studier som fås från regressionskoefficienterna (för hela modellen för sökandestocken 2018). Skillnaden syns både för de olika undermodellerna per delgrupp och som ett viktat medelvärde av dessa, som sammanfattar inskrivningstidens påverkan på populationen som helhet.

För helheten ser man att även om man försöker kontrollera för andra egenskaper framträder ett starkt samband: ökad inskrivningstid innebär ökat avstånd till arbetsmarknaden. Sambandets styrka varierar en hel del mellan olika grupper och är till och med negativt för de som är eller varit i etableringsmålgruppen. För övriga utrikesfödda är sambandet också svagare än för svenskfödda. Det har sannolikt att göra med att dessa grupper, i synnerhet etableringsgruppen, i genomsnitt etablerar sig mer i Sverige och på svensk arbetsmarknad över tid, och att inskrivningstiden är starkt korrelerad med tiden i Sverige. Det ska nämnas att tiden med uppehållstillstånd också är med som förklarande variabel, och där är sambandet negativt: det negativa sambandet mellan inskrivningstid och avstånd till arbetsmarknaden skulle förmodligen bli ännu starkare för etableringsmålgruppen om man inte kontrollerade för tiden med uppehållstillstånd.

Figur 4. Inskrivningstidens påverkan på prognosticerad arbetslöshets tid från regressionskoefficienter – alltså med ett försök att konstanthålla andra observerbara egenskaper hos de arbetssökande.³⁴



Bilaga 2. Viktningen i träffsäkerhetsmättet

Vikterna

I beräkningen av träffsäkerhetsmättet används en viktning som ger alla arbetssökande en vikt som gör att den viktade fördelningen av utfall är densamma som för den aktuella referensfördelningen (utfallen för referenspopulationen respektive de totala inskrivningstiderna för sökandestocken). Alla värden som är större än den minsta förekommande uppföljningstiden (660 dagar) betraktas som samma värde i denna viktning. Med detta i åtanke finns det i denna undersökning ett ändligt antal möjliga utfall (antalet dagar är antingen ett positivt heltal mindre än eller lika med 660 – eller så är det större än 660) vilket gör att vikten w_k för alla utfall som är lika med k_{35} blir

$$w_k = \frac{n_k^{\text{ref}}}{n_k}$$

³⁴ Beträkta till exempel två personer som är lika i alla observerbara avseenden förutom att den ena är nyinskriven och den andra har två års inskrivningstid: då bedöms den ytterligare tiden till arbete/studier i genomsnitt som drygt 20 procent längre för den senare personen. Om personen är en svenskfödd kvinna så blir ökningen 35 procent och om det är en man i etableringsmålgruppen så blir det i stället en minskning på 5 procent.

³⁵ "Lika med k " innefattar även större än 660.

där n_k är antalet utfall lika med k i den studerade populationen och n_k^{ref} är motsvarande för referensfördelningen³⁷.

Referensfördelningen

Den referensfördelning som används i denna rapport är ”medelvärdet av fördelningarna” för sökandestocken 2018 och de nyinskrivna 2018. Detta medelvärde fås genom att slå samman de två populationerna. Eftersom det är färre individer bland de nyinskrivna ges alla individer i denna population en vikt som är $N_{\text{sökandestock}}/N_{\text{nyinskrivna}}$ där $N_{\text{sökandestock}}$ är antalet individer i sökandestocken 2018 och $N_{\text{nyinskrivna}}$ är antalet individer i populationen med nyinskrivna 2018. Då fås referensfördelningen som syns i Figur 3.

Bilaga 3. Populationernas storlek

Tabell 5. Antalet arbets sökande i respektive population.

	Träningpopulation	Testpopulation
Sökandestocken 2016	37 022	36 834
Nyinskrivna 2016	28 461	28 434
Sökandestocken 2017	39 758	39 629
Nyinskrivna 2017	26 435	26 611
Sökandestocken 2018	40 729	40 624
Nyinskrivna 2018	25 914	25 792

³⁶ Även för referensfördelningen används viktning för att åstadkomma ett medelvärde mellan fördelningen för sökandestocken och de nyinskrivna, se avsnitt 10.2; ”antalet” är därför egentligen en viktad summa.

³⁷ I de fåtal fall där $n_k = 0$ hoppas k över och n_k^{ref} adderas till n_{k+1}^{ref} , med en mycket begränsad påverkan på fördelningen. Har man kontinuerliga utfall eller fler icke-förekommande värden behöver en *något* mer komplicerad viktning användas.