



Bilaga A

Statistisk analys

Ett steg in och en ny start – hur fungerar subventionerade anställningar för nyanlända? (RiR 2013:17)



Statistisk analys

Beskrivning av data och variabler

Data från Arbetsförmedlingen och SCB

Riksrevisionen har beställt data från Arbetsförmedlingen och SCB för de nyanlända som var berättigade till instegsjobb perioden 2008-2012 samt uppgifter om de instegsjobb och nystartsjobb som pågick under perioden 2008–2012. Eftersom man enligt regelverket för instegs- och nystartsjobb är nyanländ upp till tre år efter uppehållstillstånd består datasetet av personer som fått uppehållstillstånd från och med 2005 till och med 2012.

Arbetsförmedlingen har levererat data till SCB för alla nyanlända som var inskrivna under den studerade perioden. SCB har säkerställt att individer i dataseten har haft rätt skäl för uppehållstillstånd och har rensat bort de nyanlända som inte var berättigade till instegsjobb. Rensningen har gjorts utifrån Migrationsverkets klassningskoder.¹

SCB har utgått ifrån följande ”grund för bosättning”-beslut (GFB) för individerna:

- alla beslut fattade samma dag som angiven periods startdatum, fram till slutdatumet för perioden.
- finns inget beslut fattat vid startdatumet har komplettering skett med beslutet som är beviljat närmast före periodens startdatum.

SCB har exkluderat de individer ur datasetet som saknar klassningskod (i GFB-beslut) som berättigar till instegsjobb.

Efter bearbetningen av data har dataseten tillfälligt gjorts tillgängliga för Riksrevisionen i SCB:s system för mikrodataåtkomst, MONA. Innan leverans har data anonymiserats av SCB.

Variabeldefinitioner och källor

Tabell A1. Variabeldefinitioner arbetssökande (probitanalys kapitel 3 och 6)

Variabel	Definition	Källa: Arbetsförmedlingen (AF)/SCB
Kvinna	0/1-variabel lika med ett om sökande är kvinna, noll annars	Datalagret, INSPER (AF)
Ålder	Sökandes ålder i år	Datalagret, INSPER

¹ Följande klassningskoder berättigar till instegsjobb: alla koder som börjar med bokstaven A (det finns ofta antingen en bokstav till eller ett nummer efter begynnelsebokstaven), alla koder som börjar med bokstaven B, alla koder som börjar med bokstaven G samt alla koder som börjar med K, dock inte KV.

		(AF)
Barn 0-6 år	0/1-variabel lika med ett om sökande har hemmavarande barn som är 6 år eller yngre	LISA (SCB)
Utbildningsnivå	Kategorivariabel. Högst registrerad utbildningsnivå med 6 kategorier (förgymnasial utbildning < 9 år, förgymnasial utbildning > 9 år, gymnasial utbildning, eftergymnasial utbildning < två år, eftergymnasial utbildning > två år, forskarutbildning)	Datalagret, PERSONDATA (AF)
Födelseland grupperad (1)	Kategorivariabel. EU15 samt Norden, Afrika, Asien, Europa utom EU15, Nordamerika, Sydamerika och övriga.	Datalagret, INSPER (AF)
Födelseland grupperad (2)	Kategorivariabel. EU15 samt Norden, Europa utom EU15, f.d. Jugoslavien, Nordafrika 1 (Eritrea, Etiopien, Mauretanien, Sudan), Nordafrika 2 (Tunisien, Syrien, Egypten, Algeriet, Libyen, Marocko), Afghanistan, Afrika, Asien, Irak, Iran, Nordamerika, Somalia, Sydamerika, Thailand, Övriga	Datalagret, INSPER (AF)
Arbetsrelaterat handikapp	0/1-variabel lika med ett om sökande har funktionsnedsättning som medför nedsatt arbetsförmåga (enligt Arbetsförmedlingens koder för funktionshinder)	Datalagret, INSPER (AF)
Kommun	Kommun där sökande är mantalsskriven, kommunnummer.	Datalagret, INSPER (AF)
Arbetslöshetshistorik	Dagar i arbetslöshet med 4 kategorier (0 = <90 dagar, 1 = 91-180, 2=180-365 och 3=>365 dagar.)	Datalagret, SOKATPER (AF)

Tabell A2. Variabeldefinitioner arbetssökande (deskriptiv analys)

Variabel	Definition	Källa: Arbetsförmedlingen (AF)/SCB
Instegsjobbare	0/1-variabel lika med ett om arbetssökande vid något tillfälle perioden 2008–2012 haft ett instegsjobb	Övriga data, ”Program” (AF)
Nystartsjobbare	0/1-variabel lika med ett om arbetssökande vid något tillfälle perioden 2008–2012 haft ett nystartsjobb	Övriga data, ”Program” (AF)
Etableringsuppdraget	0/1-variabel lika med ett om arbetssökande vid något tillfälle perioden 2010–2012 omfattats av etableringsuppdraget	Datalagret, PERSONDATA (AF)
Tid från uppehållstillstånd till instegsjobb/nystartsjobb	Antal dagar från datum för uppehållstillstånd till startdatum av instegsjobb/nystartsjobb	Datalagret, PERSONDATA samt övriga data ”Program” (AF)

Tabell A3. Variabeldefinitioner instegsjobb och nystartsjobb (deskriptiv analys).

Variabel	Definition	Källa: Arbetsförmedlingen (AF)/SCB
Bransch	Kategorivariabel enligt indelning för SNI2007	Övriga data, ”Program” (AF)
Yrke	Kategorivariabel enligt indelning för SSK 96	Övriga data, ”Program” (AF)
Omfattning	Procentuell tjänstgöringsgrad	Övriga data, ”Program” (AF)
Lön	Kronor, omräknat till månadslön (40 h/v)	Övriga data, ”Program” (AF)
Storlek företag	Antal sysselsatta	LISA (SCB)
Sektor	Kategorivariabel med 10 kategorier (Statlig förvaltning, Statliga affärsverk, Primärkommunal förvaltning, Landsting, Övriga offentliga institutioner, Aktiebolag ej offentligt ägda, Övriga företag ej offentligt ägda, Statligt ägda företag och organisationer, Kommunalt ägda företag och organisationer, Övriga organisationer)	LISA (SCB)

Beskrivande statistik

Berättigade till instegs- och nystartsjobb 2012.

Tabell A4. Könsfördelning. Antal och procent.

Kön	Antal	Procent
Män	31 146	49,72
Kvinnor	31 501	50,28
<i>Totalt</i>	<i>62 647</i>	<i>100,00</i>

Tabell A5. Åldersfördelning. Medelvärde och standardavvikelser.

Medelvärde	Standard avvikelse	Min	Max
33,47	9,29	20	66

Tabell A6. Födelseland, grupperad (1). Procentuell fördelning.

EU15 samt Norden	0,13
Afrika	32,99
Asien	46,22
Europa utom EU15	14,43
Nordamerika	2,25
Sydamerika	3,10
Övriga	0,90
<i>Totalt</i>	<i>100,00</i>

Not: används i probitanalys kapitel 3.

Tabell A7. Födelseland, grupperad (2). Procentuell fördelning.

EU15 samt Norden	0,13
Fd Jugoslavien	7,55
Nordafrika 1	8,51
Nordafrika 2	9,68
Afghanistan	5,08
Afrika	5,65
Asien	12,31
Europa utom EU15	6,97
Irak	12,63
Iran	4,85
Nordamerika	2,25
Somalia	15,17
Sydamerika	3,10
Thailand	5,23
Övriga	0,90
<i>Totalt</i>	<i>100,00</i>

Not: används i probitanalys kapitel 6 samt deskriptiv analys

Definitioner av arbete

Tabell A8. Definitioner av arbete efter instegs- och nystartsjobb (kapitel 3).

Definition	Omfattar
Arbete 1	Osubventionerade heltidsanställningar på den öppna arbetsmarknaden Reguljär utbildning
Arbete 2	Arbete 1+ Deltidsanställningar Timanställningar Tillfälligt arbete
Arbete 3 (avser enbart instegsjobb)	Arbete 2+ Lönebidrags-anställningar Offentligt skyddade anställningar Nystartsjobb

Not: Definitionerna utgår från de avaktualiseringskoder och sökandekategorier som finns i Arbetsförmedlingens statistik. Se AMV (2007).

Arbete 1 utgör den snävaste definitionen av arbete; här räknas endast osubventionerade heltidsanställningar på den öppna arbetsmarknaden samt övergång till reguljär utbildning som lyckade arbetsmarknadsutfall.² Många nyanlända står dock relativt långt från den ordinarie arbetsmarknaden. Därför omfattar *arbete 2* alla arbetsmarknadsutfall i *arbete 1* samt deltidsanställningar, timanställningar eller tillfälliga arbeten.³ I analysen av dem som haft instegsjobb har vi slutligen inkluderat övergångar från instegsjobb till nystartsjobb i den tredje definitionen (*arbete 3* avser således endast analysen av instegsjobb).⁴ I definitionen ingår även lönebidragsanställningar och offentligt skyddade anställningar.⁵ Dessa anställningsformer är främst till för personer med dokumenterade arbetshandikapp; andelen arbetssökande som gick sådana anställningar var vidare endast ett fåtal och påverkar endast resultatet marginellt. Skillnaden mellan *arbete 2* och *arbete 3* kan därför ses som att definitionen endast utökas till att även omfatta nystartsjobb.

² I Arbetsförmedlingens statistik är detta en övergång från sökandekategori 78 (instegsjobb) eller 33 (nystartsjobb) till är detta avaktualiseringsorsak 1 (Tillsvidareanställning eller egen näringsverksamhet utan stöd), 2 (Tidsbegränsad anställning), 3 (Fortsatt anställning hos samma arbetsgivare), 4 (Anställning inom Samhall) samt 7 (Annan utbildning än arbetsmarknadsutbildning).

Avaktualiseringsorsak 4 innebär i och för sig en anställning på samhall och är en subventionerad anställning, dock inte inom ramen för arbetsmarknadspolitiken.

³ Sökandekategori 21 (deltidsanställd) eller 22 (timanställd) eller 31 (tillfälligt arbete). En väsentlig skillnad är att dessa individer fortfarande kan vara registrerade som arbetssökande.

⁴ Sökandekategori 33.

⁵ Sökandekategori 42 respektive 43.

Uppföljningsperiod och gruppering av data

Detta avsnitt redogör för de definitioner och avgränsningar i data som har gjorts för överlevnads- och probitanalyser i kapitel 3.

Population och uppföljningsperiod

Populationen utgörs av de personer som var berättigade till instegsjobb eller nystartsjobb för nyanlända någon gång under 2009. Uppföljningsperioden är vald för att hålla villkoren för den berättigade populationen konstant. Eftersom man enligt regelverket för instegs- och nystartsjobb är nyanländ upp till tre år efter uppehållstillstånd består den studerade populationen av personer som fått uppehållstillstånd åren 2006–2009.⁶

Individerna i populationens arbetsmarknadsstatus har följts under perioden 2009–2012 via de *sökandekategorier* och *avaktualiseringsorsaker* som registrerats i Arbetsförmedlingens datalager.⁷ De arbetssökande delas in i olika sökandekategorier, exempelvis öppet arbetslös, i instegsjobb, i nystartsjobb etc. När någon lämnar Arbetsförmedlingen görs en kodad notering om orsaken, exempelvis att personen fått ett ordinarie arbete, fått ett vikariat, avlidit, okänd orsak etc.⁸ Dessa uppgifter har i Riksrevisionens undersökning kompletterats med andra individuppgifter från SCB (se tidigare avsnitt). Riksrevisionens observerade arbetsmarknadsutfall för de nyanlända har skett vid den *första* tidpunkt som de av någon orsak har lämnat Arbetsförmedlingen. Detta innebär att om individer gått till arbete någon gång 2009–2012 och blivit arbetslösa vid ett senare tillfälle under perioden 2009–2012 räknas dessa i analysen som att de har fått arbete.

Gruppering av individer

Nedan beskriver vi kriterier för grupperna instegsjobbare respektive nystartsjobbare samt jämförelsegruppen.

Inom nationalekonomin brukar man tala i termer av ”behandling” när man ska utvärdera programeffekter av arbetsmarknadspolitiska åtgärder. Även om vi inte kan utvärdera de rena programeffekterna för instegsjobb och nystartsjobb använder vi ändå termen behandlad när vi jämför arbetsmarknadsresultaten för olika grupper. Stödjobbarna ses som ”behandlade” grupper och dessa ska ha fått sin behandling under 2009, inte före 2009. Nystartsjobbarna kan dock ha haft instegsjobb före sitt nystartsjobb.⁹ Jämförelsegruppen ska dock inte ha fått någon behandling (i analysen av *arbete3* där nystartsjobb utgör ett utfall snarare än en behandling

⁶ Det finns dock vissa individer som har varit inskrivna vid Arbetsförmedlingen tidigare än 2006, vilket kan röra sig om individer som har flera beslut för uppehållstillstånd med olika kvalificeringsgrunder. Data har dock begränsats till att endast avse individer som skrevs in hos Arbetsförmedlingen första gången under eller efter 2005.

⁷ Se AMV (2007).

⁸ När en individ avaktualiseras för första gången efter 2009 lämnar personen Riksrevisionens undersökning även om den senare under perioden återinskrivs.

⁹ Instegsjobbet i sig utgör därmed en sorts ”tidigare behandling” men eftersom vi särskilt studerar de som tidigare har haft instegsjobb och jämför med de som bara har haft nystartsjobb isolerar vi påverkan av denna tidigare behandling.

tillåts jämförelsegruppen ha haft nystartsjobb under eller efter 2009 eftersom instegsjobbarna kan ha hunnit få ett nystartsjobb).

ARBETE 1–2

Instegsjobbare är individer som uppfyller följande:

- De *ska* ha påbörjat (minst) ett instegsjobb under 2009 (grundkriteriet)
- De *får varken* ha haft instegs- eller nystartsjobb före 2009
- De *kan* ha haft nystartsjobb under 2009
- De *kan* ha haft instegs- och/eller nystartsjobb efter 2009 (dvs. 2010–2012).

Nystartsjobbare

- De *ska* ha påbörjat (minst) ett nystartsjobb under 2009 (grundkriteriet)
- De *får inte* ha haft nystartsjobb före 2009
- De *kan* ha haft instegsjobb före, under och efter 2009 (dvs. 2007–2012)
- de *kan* ha haft nystartsjobb efter 2009 (dvs. 2010–2012).

Jämförelsegrupp

- De *får inte* ha haft instegs- eller nystartsjobb vare sig före, efter eller under 2009 (grundkriteriet). Det vill säga att jämförelsegruppen inte har haft något av stöden under perioden 2007–2012.

Tabell A9. Matris, definition av grupper för probitanalys kapitel 3, arbete 1–2.

Arbete 1–2	2007–2008		2009		2010–2012	
	INS	NYS	INS	NYS	INS	NYS
Instegsjobbare			<u>x</u>			
Nystartsjobbare				<u>x</u>		
Jämförelsegrupp			<u>x</u>	<u>x</u>		

Not: Grön ruta innebär att gruppen *kan* ha haft stödet (men det är inget *krav*) och ett understruket kryss utgör grundkriteriet för gruppen. Röda rutor innebär att gruppen ej har haft stödet under angiven tidsperiod.

ARBETE 3

Följande gäller bara analys av instegsjobben. Gruppen instegsjobbare är densamma som i arbete 1–3, däremot har jämförelsegruppen utvidgats.

Jämförelsegrupp

- De *får inte* ha haft instegsjobb under 2009 (grundkriteriet)
- De *får inte* ha haft instegs- eller nystartsjobb före 2009 (2007–2008).
- De *får inte* ha haft instegsjobb perioden efter 2009 (2010–2012).
- De *kan* ha haft nystartsjobb under 2009.

Tabell A10. Matris, definition av grupper för probitanalys kapitel 3, arbete 3.

Arbete 3	2007–2008		2009		2010–2012	
	INS	NYS	INS	NYS	INS	NYS
Instegsjobbare			<u>x</u>			
Jämförelsegrupp			<u>x</u>			

Not: Grön ruta innebär att gruppen *kan* ha haft stödet (men det är inget *krav*) och ett understruket kryss utgör grundkriteriet för gruppen. Röda rutor innebär att gruppen ej har haft stödet under angiven tidsperiod.

Resultat probitanalys

Probitanalys i kapitel 3

Tabell A11. Sannolikheten att få ett arbete enligt definition 1, 2 och 3 fram till 31/12 2012. Instegsjobb. Marginaleffekter.

Utfallsdefinition	Arbete 1	Arbete 2	Arbete 3
Instegsjobb	-0,067	-0,011	0,179
Ålder	0,002	0,009	0,018
Ålder i kvadrat	-0,0001	-0,0002	-0,0003
År i Sverige	0,038	0,046	0,038
Kvinna	-0,098	-0,090	-0,121
Barn 0–6	-0,061	-0,067	-0,088
Arbetsrelaterat handikapp	-0,361	-0,358	-0,177
<i>Födelseland grupperat</i>			
(EU15 samt Norden referens)			
Afrika	-0,090	-0,011	-0,045
Asien	-0,109	-0,059	-0,051
Europa utom EU15	-0,030	0,047	0,033
Nordamerika	0,016	0,077	0,003
Sydamerika	0,049	0,129	0,068
Övriga	0,016	0,102	0,015
<i>Högsta utbildning</i>			
Gymnasium (Grundskola ref)	0,081	0,082	0,076
Eftergymnasialutbildning	0,119	0,119	0,113
<i>Arbetslöshetshistorik</i>			
(<90 dagar referens)			
91–180 dagar	-0,027	-0,024	-0,026
180–365 dagar	-0,057	-0,067	-0,052
>365 dagar	-0,102	-0,109	-0,164
Pseudo R2	0,106	0,119	0,132

Fet stil = signifikant på 5-procentsnivån, *kursiv stil* = signifikant på 10-procentsnivån

Not: Dummyvariabler för kommun har ingått i modellen, marginaleffekterna redovisas ej.

Tabell A12. Sannolikheten att få ett arbete enligt definition 1 och 2 fram till 31/12 2012. Nystartsjobb med föregående instegsjobb. Marginaleffekter.

Utfallsdefinition	Arbete 1	Arbete 2
Nystartsjobb	0,063	0,151
Ålder	0,005	0,012
Ålder i kvadrat	-0,0002	-0,0003
År i Sverige	0,040	0,048
Kvinna	-0,098	-0,088
Barn 0–6	-0,066	-0,071
Arbetsrelaterat handikapp	-0,366	-0,358
<i>Födelseland grupperat</i> (EU15 samt Norden referens)		
Afrika	-0,059	0,003
Asien	-0,084	-0,054
Europa utom EU15	-0,009	0,049
Nordamerika	0,034	0,081
Sydamerika	0,058	0,126
Övriga	0,016	0,100
<i>Högsta utbildning</i>		
Gymnasium	0,089	0,090
Eftergymnasial utbildning	0,127	0,131
<i>Arbetslöshetshistorik</i> (<90 dagar referens)		
91-180 dagar	-0,040	-0,035
180-365 dagar	-0,075	-0,083
>365 dagar	-0,118	-0,126
Pseudo R2	0,114	0,131

Fet stil = signifikant på 5-procentsnivån, *kursiv stil* = signifikant på 10-procentsnivån.

Not: Dummyvariabler för kommun har ingått i modellen, marginaleffekterna redovisas ej.

Tabell A13. Sannolikheten att få ett arbete enligt definition 1 och 2 fram till 31/12 2012. Nystartsjobb utan föregående instegsjobb. Marginal effekter.

Utfallsdefinition	Arbete 1	Arbete 2
Nystartsjobb	0,020	0,077
Ålder	0,002	0,009
Ålder i kvadrat	-0,0001	-0,0002
År i Sverige	0,038	0,046
Kvinna	-0,101	-0,090
Barn 0–6	-0,055	-0,057
Arbetsrelaterat handikapp	-0,360	-0,357
<i>Födelseland grupperat</i> (EU15 samt Norden referens)		
Afrika	-0,078	-0,049
Asien	-0,105	-0,103
Europa utom EU15	-0,028	-0,003
Nordamerika	0,026	0,052
Sydamerika	0,041	0,086
Övriga	0,015	0,040
<i>Högsta utbildning</i>		
Gymnasium	0,089	0,089
Eftergymnasial utbildning	0,127	0,126
<i>Arbetslöshetshistorik</i> (<90 dagar referens)		
91–180 dagar	-0,028	-0,025
180–365 dagar	-0,065	-0,069
>365 dagar	-0,107	0,116
Pseudo R2	0,110	0,125

Fet stil = signifikant på 5-procentsnivån, *kursiv stil* = signifikant på 10-procentsnivån.

Not: Dummyvariabler för kommun har ingått i modellen, marginaeffekterna redovisas ej.

Tabell A14. Sannolikheten att få ett arbete enligt definition 1 och 2 fram till 31/12 2012. Nystartsjobb med föregående instegsjobb (nystartsjobbare utan föregående instegsjobb är jämförelsegrupp). Marginaleffekter.

Utfallsdefinition	Arbete 1	Arbete 2
Instegsjobb före nys	0,049	0,076
Ålder	-0,020	-0,009
Ålder i kvadrat	0,0001	-0,00002
År i Sverige	0,008	0,009
Kvinna	-0,062	-0,049
Barn 0–6	-0,010	-0,058
Arbetsrelaterat handikapp	-0,354	-0,266
<i>Födelseland grupperat</i> (EU15 samt Norden referens)		
Afrika	-0,290	-0,950*
Asien	-0,252	-0,902*
Europa utom EU15	-0,169	-0,967*
Nordamerika	0,024	-0,840*
Sydamerika	0,082	-0,862*
Övriga	-0,348	-0,840*
<i>Högsta utbildning</i>		
Gymnasium	0,118	0,065
Eftergymnasial utbildning	0,141	0,085
<i>Arbetslöshetshistorik</i> (<90 dagar referens)		
91–180 dagar	-0,037	-0,009
180–365 dagar	-0,061	-0,023
>365 dagar	-0,013	-0,066
Pseudo R2	0,107	0,124

Fet stil = signifikant på 5-procentsnivån, *kursiv stil* = signifikant på 10-procentsnivån.

* De höga marginaleffekterna visar på att vi har för få i varje födelselandskategori. Vi har skattat en modell med färre kategorier; övriga skattningar (förutom de för födelselandvariablerna) förändrades endast marginellt.

Not: Dummyvariabler för kommun har ingått i modellen, marginaleffekterna redovisas ej.

Probitanalys i kapitel 6

Tabell A15-16 visar två probitskattningar med olika populationer, resultatet för variabeln "Kvinna" ändras dock bara marginellt.

Tabell A15. Sannolikheten att få ett instegsjobb under 2011, population: berättigade 2011. Marginaleffekter.

	Marginaleffekt
Ålder	0,007
Ålder i kvadrat	-0,0001
Kvinna	-0,046
År i Sverige	-0,00007
Barn 0–6	-0,024
Arbetsrelaterat handikapp	-0,026
I etableringsuppdraget	-0,077
<i>Födelseland grupperat</i>	
(EU15 samt Norden referens)	
F.d. Jugoslavien	0,027
Nordafrika (1)	0,022
Nordafrika (2)	0,031
Afghanistan	-0,001
Afrika	0,00004
Asien	0,019
Europa utom EU 15 och Norden	0,015
Irak	0,012
Iran	0,030
Nordamerika	0,023
Somalia	0,005
Sydamerika	0,014
Thailand	0,075
Övriga	0,023
<i>Högsta utbildning</i>	
(förgymnasial utbildning < 9 år referens)	
Förgymnasial utbildning > 10 år	0,002
Gymnasial utbildning	0,007
Eftergymnasial utbildning < 2 år	-0,002
Eftergymnasial utbildning > 2 år	-0,013
Forskarutbildning	-0,044
Pseudo R2	0,096

Fet stil = signifikant på 5-procentsnivån (ingen variabel signifikant på 10-procentsnivån i denna skattning)

Not: Dummyvariabler för kommun har ingått i modellen, marginaleffekterna redovisas ej.

Tabell A16. Sannolikheten att få ett instegsjobb under 2012, population: berättigade 2012 som fick uppehållstillstånd efter 1 dec 2010 (då etableringsreformen infördes). Marginaleffekter.

	Marginaleffekt
Ålder	0,001
Ålder i kvadrat	-0,00004
Kvinna	-0,057
År i Sverige	0,035
Arbetsrelaterat handikapp	-0,022
I etableringsuppdraget	-0,035
<i>Födelseland grupperat</i> (EU15 samt Norden referens)	
F.d. Jugoslavien	0,047
Nordafrika (1)	0,026
Nordafrika (2)	0,031
Afghanistan	0,004
Afrika	0,021
Asien	0,040
Europa utom EU 15 och norden	0,026
Irak	0,015
Iran	0,050
Nordamerika	0,031
Somalia	0,016
Sydamerika	0,058
Thailand	0,142
Övriga	0,007
<i>Högsta utbildning</i> (förgymnasial utbildning < 9 år referens)	
Förgymnasial utbildning > 10 år	0,001
Gymnasial utbildning	0,008
Eftergymnasial utbildning < 2 år	0,010
Eftergymnasial utbildning > 2 år	-0,004
Forskarutbildning	-0,025
Pseudo R2	0,115

Fet stil = signifikant på 5-procentsnivån, *kursiv stil* = signifikant på 10-procentsnivån.
Not: Dummyvariabler för kommun har ingått i modellen, marginaleffekterna redovisas ej.

Tabell A17. Sannolikheten att få ett inestegsjobb 2011, population: kvinnor respektive män av berättigade 2011. Marginaleffekter.

	Marginaleffekt	
	Kvinnor	Män
Ålder	0,005	0,009
Ålder i kvadrat	-0,00007	-0,0001
År i Sverige	-0,00003	-0,0001
Barn 0–6	-0,034	-0,010
Arbetsrelaterat handikapp	-0,013	-0,038
I etableringsuppdraget	-0,059	-0,100
<i>Födelseland grupperat</i> (EU15 samt Norden referens)		
Fd Jugoslavien	-0,009	0,081
Nordafrika (1)	0,002	0,050
Nordafrika (2)	-0,010	0,091
Afghanistan	-0,022	0,032
Afrika	-0,018	0,032
Asien	-0,003	0,045
Europa utom EU 15 och Norden	-0,021	0,070
Irak	-0,022	0,064
Iran	-0,002	0,079
Nordamerika	0,003	0,047
Somalia	-0,018	0,045
Sydamerika	-0,011	0,051
Thailand	0,025	<i>0,144</i>
Övriga	-0,0005	0,062
<i>Högsta utbildning</i> (förgymnasial utbildning < 9 år referens)		
Förgymnasial utbildning > 10 år	0,003	-0,0006
Gymnasial utbildning	0,007	0,007
Eftergymnasial utbildning < 2 år	0,001	-0,005
Eftergymnasial utbildning > 2 år	-0,006	-0,022
Forskarutbildning	-0,040	-0,048
Pseudo R2	0,100	0,094

Fet stil = signifikant på 5-procentsnivån, *kursiv stil* = signifikant på 10-procentsnivån.

Dummyvariabler för kommun har ingått i modellen, marginaleffekterna redovisas ej.

Tabell A18. Sannolikheten att få ett inestegsjobb 2011, population: berättigade 2011 med barn 0–6 år. Marginaleffekter.

	Marginaleffekt
Ålder	0,002
Ålder i kvadrat	-0,00004
År i Sverige	-0,00006
Kvinna	-0,062
Arbetsrelaterat handikapp	-0,022
I etableringsuppdraget	-0,060
<i>Födelseland grupperat*</i> (Europa och Nordamerika referens)	
Afrika	-0,005
Asien	-0,004
Sydamerika	-0,017
Övriga	-0,002
<i>Högsta utbildning</i> (förgymnasial utbildning < 9 år referens)	
Förgymnasial utbildning > 10 år	0,002
Gymnasial utbildning	0,005
Eftergymnasial utbildning < 2 år	0,002
Eftergymnasial utbildning > 2 år	-0,011
Forskarutbildning	-0,040
Pseudo R2	0,1294

Fet stil = signifikant på 5 procentsnivån, *kursiv stil* = signifikant på 10 procentsnivån.

Dummyvariabler för kommun har ingått i modellen, marginaleffekterna redovisas ej.

*För få observationer i flera grupper krävde en omgruppering av födelselandsvariabeln med färre kategorier

Teknisk beskrivning av metoder

Probitmodellen

I kapitel 3 och 6 används en probitmodell. I kapitel 3 är syftet att analysera sambandet mellan deltagande i instegsjobb respektive nystartsjobb och sannolikheten att få ett osubventionerat arbete. I kapitel 6 används modellen för att undersöka sambandet mellan kön och sannolikheten att få ett instegsjobb. Probitmodellen används då den beroende variabeln i dessa båda fall är binär, det vill säga den kan bara anta två värden (ha fått arbete/inte fått arbete respektive ha fått instegsjobb/inte fått instegsjobb).

Probitmodellen definieras som:

$$\Pr(Y = 1|X) = \Phi(X'\beta)$$

Där Y är vår beroende variabel. \Pr betecknar sannolikheten och Φ är den kumulativa fördelningsfunktionen för normalfördelningen. X är en vektor av oberoende variabler som antas påverka utfallet för Y . Parametrarna β skattas genom s.k. maximum likelihood-estimation.¹⁰

De redovisade skattningarna i tabell A11–18 visar på marginaleffekten av olika variabler på den beroende variabeln Y (arbete i tabell A11–14 samt instegsjobb i tabell A15–18). Exempelvis är den skattade marginaleffekten för variabeln ”Kvinna” -0,046 i tabell A15. Detta värde tolkas som att det är i genomsnitt 5 procentenheters lägre sannolikhet för en kvinna att få ett instegsjobb jämfört med en man, när hänsyn har tagits till skillnader i övriga variabler i modellen.

Överlevnadskurvor och Kaplan-Meier-metoden

I kapitel 3 skattas överlevnadskurvor med Kaplan-Meier-metoden. Inom medicinsk forskning används Kaplan-Meier-metoden för att jämföra överlevnadsfrekvensen för en grupp som har fått behandling med överlevnadsfrekvensen för en grupp som inte har fått behandling. Inom nationalekonomin används bl.a. metoden för att jämföra tiden till arbete för två grupper av arbetslösa, där den ena gruppen har fått ”behandling” i form av någon arbetsmarknadspolitisk åtgärd. I den här granskningen jämförs instegsjobbare respektive nystartsjobbare med arbetslösa som inte har tagit del av något utav stöden.

¹⁰ Se även Gujarati och Porter (2009).

Anta att det finns k skilda tidpunkter, $t_1 < t_2 < \dots < t_k$. Vid varje tidpunkt t_j finns det n_j individer som har en chans att få ett arbete. Med chans menas att de inte före tiden t_j har fått ett arbete eller har försvunnit från arbetsförmedlingen av okänd orsak¹¹ (censurering). Låt d_j vara antal individer som får ett arbete vid tidpunkten t_j . Kaplan-Meier-estimatoren definieras då som:

$$\hat{S}(t) = \prod_{j:t_j \leq t} \left[1 - \frac{d_j}{n_j} \right]$$

Överlevnadsfunktionen $S(t_j)$ ger därmed sannolikheten att vara arbetslös vid tidpunkten t_j .¹² Eftersom S är en sannolikhet vet vi att funktionen bara kan anta värden mellan 0 och 1. Vidare vet vi att eftersom tiden T aldrig kan vara negativ ger det att $S(0)=1$. Inom dessa restriktioner kan sannolikhetskurvorna anta olika former men har alltid negativ eller ingen lutning.

$S(t_j)$ beräknas enligt följande, för varje given tidpunkt t : 1) beräkna för varje tidigare tidpunkter summan inom parentes, vilket kan tolkas som den beroende sannolikheten att vara arbetslös till tidpunkten t_{j+1} givet överlevnad till tidpunkten t_j . 2) Därefter multipliceras alla dessa beroende sannolikheter.¹³

I kapiteltexten har vi vidare valt att invertera kurvan för att tolkningen ska bli *sannolikheten att ha fått arbete* vid tidpunkten t (istället för sannolikheten att vara arbetslös).

¹¹ Avaktualiseringskod 5 i Arbetsförmedlingens statistik.

¹² Överlevnadsfunktionen $S(t_j)$ ger formellt sett sannolikheten att vara arbetslös till och med t_j och fram till tidpunkten (men inte till och med) t_{j+1} . I vår analys kan dock överlevnadskurvorna approximativt tolkas som sannolikheten att vara arbetslös vid t_j eftersom datasetet är relativt stort och tidpunkterna därför kan ses som kontinuerliga (med andra ord är tiden från t_j till t_{j+1} relativt kort och sannolikheten att få ett arbete under den här tiden kan bortses ifrån). Denna tolkning av överlevnadskurvorna är mer intuitiv och används därför för enkelhets skull i kapitel 3.

¹³ Se även Allison (2007) och Cleves m.fl. (2010) för mer om Kaplan-Meier metoden.