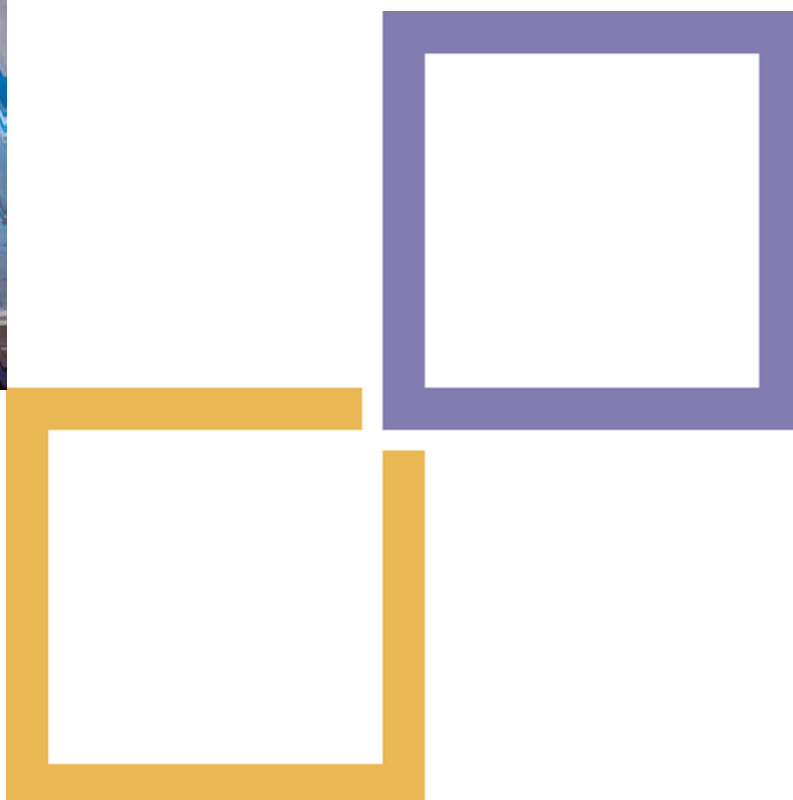


Henrik Lindegaard Andersen

Forklarende analyse af ikke-vestlige indvandreres arbejdsmarkedstilknnytning



Publikationen *Forklarende analyse af ikke-vestlige indvandreres arbejdsmarkedstilknytning* kan downloades fra hjemmesiden www.kora.dk

© KORA og forfatteren

Mindre uddrag, herunder figurer, tabeller og citater, er tilladt med tydelig kildeangivelse. Skrifter, der omtaler, anmelder, citerer eller henviser til nærværende, bedes sendt til KORA.

© Omslag: Mega Design og Monokrom

Udgiver: KORA

ISBN: 978-87-7509-501-8

5244

December 2012

KORA

Det Nationale Institut for

Kommuners og Regioners Analyse og Forskning

KORA er en uafhængig statslig institution, hvis formål er at fremme kvalitetsudvikling, bedre ressourceanvendelse og styring i den offentlige sektor.



**Det Nationale Institut
for Kommuners og Regioners
Analyse og Forskning**

Købmagergade 22
1150 København K
E-mail: kora@kora.dk
Telefon: 444 555 00

Henrik Lindegaard Andersen

Forklarende analyse af ikke-vestlige indvandreres arbejdsmarkedstilknnytning

KORA, Det Nationale Institut
for Kommuner og Regioners Analyse og Forskning
2012

Forord

Nærværende rapport undersøger determinanter for ikke-vestlige indvandrere, efterkommere samt danskeres tilknytning til arbejdsmarkedet. Rapporten er udarbejdet i sensommeren 2012 af KORA – Det Nationale Institut for Kommuners og Regioners Analyse og Forskning – for Social- og Integrationsministeriet. Forsker Henrik Lindegaard Andersen, der har udarbejdet analyserne, ønsker at takke følgegruppen, som har bestået af medarbejdere fra Social- og Integrationsministeriet samt repræsentanter for Arbejdsmarkedsstyrelsen og Ankestyrelsen, for deres kommentarer. Ligeledes takkes to uafhængige lektører for deres kommentarer til tidligere udkast af rapporten.

Henrik Lindegaard Andersen
December 2012

Indhold

Sammenfatning	7
1 Baggrund	10
2 Data og metode	12
3 Beskrivende statistik.....	17
4 Analyseresultater	27
4.1 Modeller for lønindkomst	28
4.2 Udvidet model for lønindkomst over 125.000 kr.	36
4.3 Udvidet model for ingen lønindkomst	46
4.4 Udvidet model for lønindkomst over 120.000 kr. i 2008	52
4.5 Model for overgang til førtidspension i 2007–2012.....	54
Litteratur	61
Bilag 1: Konstruktion af sprogindeks	62
Bilag 2: Separat analyse for selvstændiges arbejdsmarkedstilknnytning	66
Bilag 3: Alternative modelspecifikationer	69
Bilag 4: Løn, selvforsørgelse og førtidspension blandt ikke-vestlige indvandrere og sammenlignelige danskere	73
English Summary	80

Sammenfatning

Denne rapport er blevet til på opfordring fra Social- og Integrationsministeriet. Motivationen for undersøgelsen er ønsket om en samlet analyse af, hvilke karakteristika ved personen der er bestemmende for, i hvor høj grad indvandrere med ikke-vestlig herkomst og deres efterkommere deltager på arbejdsmarkedet. Interessen for emnet udspringer også af den lavere andel af indvandrere i beskæftigelse og en højere andel, der modtager overførselsindkomster.

På baggrund af det seneste, og os bekendt største, empiriske materiale præsenterer vi i rapporten en statistisk analyse, der giver et billede af, hvordan personlige karakteristika påvirker sandsynligheden for at have en vis minimumsløn, for at være afhængig af offentlig eller familiens forsørgelse samt risikoen for at gå på førtidspension. Vores liste af individuelle karakteristika er omfattende. Vi inkluderer ikke blot oplysninger om personens køn, alder, bopæl, civilstand, oprindelse, opholdsgrundlag, erhvervs erfaring, uddannelse og sprog, men også om somatiske og mentale helbredsforhold, der kan være vigtige for, hvordan vedkommende klarer sig på arbejdsmarkedet.

Fokus i denne rapport ligger på individets karakteristika og ikke på den samfundsmæssige indretning. En analyse af denne type kan, under visse forudsætninger, bruges i det politiske arbejde til at udpege særlige områder, hvor der er grund til at iværksætte en ny – eller styrke en eksisterende – indsats. Imidlertid er det vigtigt at huske, at selv en stærk statistisk sammenhæng ikke nødvendigvis beviser en faktisk sammenhæng. Nærværende rapport påviser kun statistiske sammenhænge, og politiske anbefalinger bør derfor ikke alene hvile på resultaterne præsenteret heri.

Rapporten har rent beskrivende vist, at de ikke-vestlige indvandrere i Danmark adskiller sig på en række punkter ud over den svagere arbejdsmarkedstilknnytning. Indvandrerne er yngre, de er oftere gift, de har flere børn, og de har oftere kun en grundskoleuddannelse. På nogle områder er deres helbred lidt bedre, mens det på andre områder er lidt værre, men generelt er det overraskende, hvor høj en andel af både danskere, indvandrere og efterkommere, der lider af behandlingskrævende sygdomme. Med hensyn til arbejdsmarkedstilknnytning, kan det indvendes, at de nævnte forhold netop kan være forklaringen på den lavere beskæftigelse blandt indvandrerne. Imidlertid forholder det sig sådan, at hvis man sammenligner de ikke-vestlige indvandreres løn, selvforsørgelse og overgang til førtidspension med en gruppe af danskere, der er demografisk, uddannelses- og helbredsmæssigt identiske, så mindskes forskellen ikke væsentligt.

I forhold til den registerbaserede statistiske analyse af, hvilke faktorer der forklarer tilknytningen til arbejdsmarkedet, så har vi set, at sandsynligheden for at have en rimelig minimumsløn var lavere for helt unge og ældre indvandrere, mens personer med en erhvervsfaglig eller videregående uddannelse klarede sig bedre. Det er overraskende at finde, at sprog og opholdsgrundlaget ikke betyder noget for indvandreres løn, når vi samtidig tager højde for personens arbejdsmarkedsstatus. Nuværende og historisk god tilknytning til arbejdsmarkedet er i sig selv særdeles vigtigt for den fremtidige tilknytning til arbejdsmarkedet. Justeres der ikke for arbejdsmarkedsstatus, så har personer med sprogvanskeligheder og flygtninge

det markant vanskeligere på arbejdsmarkedet. Vi fandt også, at det påvirkede sandsynligheden negativt for indvandrere, hvis de havde oprindelse i Mellemøsten. Endelig var der stærke indikationer for, at dårligt psykisk helbred er en vigtig risikofaktor i arbejdsmarkedssammenhæng.

I det store og hele er de faktorer, der øger risikoen for at tjene under et vist minimum, også de faktorer, der øger risikoen for, at en person får overførselsindkomst eller forsørges af familien. Ældre, kvinder og ugifte har højere risiko for ikke at være lønmodtagere, men også lav uddannelse, høj ledighed eller ringe erhvervserfaring er stærke risikofaktorer. Kort opholdstid i Danmark samt oprindelse i Mellemøsten var også her risikofaktorer specielt for kvinder, mens mænd fra visse asiatiske lande havde højere chance for at være lønmodtagere. Vi kunne også konstatere, at det for kvinder øgede risikoen, hvis de var flygtede hertil eller havde sprogproblemer, hvorimod det for mænd ikke havde negativ virkning. Med hensyn til helbredet, så var det specielt psykisk sygdom, der slog negativt ud i vores undersøgelse, men også reumatiske sygdomme havde en betydelig negativ virkning.

Med hensyn til førtidspension fandt vi, at en højere alder var en vigtig risikofaktor, ligesom at mænd har en højere risiko for at få førtidspension. Højere uddannelse og større erhvervsgrad sænkede derimod risikoen. Indvandrere, der har opholdt sig i Danmark i mange år, havde også en langt højere risiko for at forlade arbejdsmarkedet tidligt. Igen var det overvejende personer fra Mellemøsten, der havde højere risiko, mens personer fra Asien havde lavere risiko. For personer med flygtningebaggrund var risikoen højere, mens personer uden sprogproblemer havde lavere risiko. Endelig så vi, at den absolut største risikofaktor, som forventet, var psykisk sygdom efterfulgt af reumatiske sygdomme.

Det mønster, som vores analyser har tegnet, er ganske klart. Hvis vi et øjeblik ser bort fra de metodiske forbehold, der er forbundet med en korrelationsanalyse, viser vores resultater, at man ville kunne øge andelen af ikke-vestlige indvandrere, der har en god tilknytning til arbejdsmarkedet, hvis man kunne fjerne de barrierer, der er forbundet med psykisk sygdom. Der er to væsentlige problemer i forhold til denne konklusion. For det første ved vi ikke, om korrelationen udtrykker en faktisk årsagssammenhæng. For det andet ved vi ikke, hvilke mekanismer der gør, at personer, der behandles for en psykisk sygdom, klarer sig dårligere på arbejdsmarkedet. En analyse, der retter sig specifikt mod disse spørgsmål, kan derfor vise sig at være værdifuld.

Et andet forhold, der springer i øjnene, er, at personer født i Mellemøsten klarer sig markant dårligere, selv efter at vi har taget højde for sproglige barrierer, uddannelse og opholdsgrundlag. Det forhold behøver ikke nødvendigvis at skyldes forhold ved individet, men kan være en konsekvens af samfundsmæssige barrierer. Vi kender ikke årsagen, men eksempler kan inkludere diskrimination eller manglende anerkendelse af indvandrerens kvalifikationer.

Det er en kendt sag, at uddannelse er vigtigt for, hvordan man klarer sig i arbejdslivet, og det viser vores undersøgelse også. Hvis en indvandrer har fuldført en erhvervsfaglig eller videregående uddannelse, så er der ikke blot større chance for, at han eller hun deltager aktivt på arbejdsmarkedet, men også langt lavere risiko for, at vedkommende modtager overførselsindkomst.

Det sidste forhold, vi vil fremhæve, er, hvad de sproglige kompetencer betyder for tilknytningen til arbejdsmarkedet. For kvinder viste vores resultater, at bedre sprogkompetencer mindsker risikoen for at modtage overførselsindkomst. Men når vi så på lønmodtagerne generelt, var sproget helt uden betydning, hvis vi samtidig justerede for stilling og erhvervs erfaring. Derimod var sprog meget vigtigt, hvis vi ikke tog hensyn til stilling og erhvervs erfaring. En intuitiv forklaring kan være, at erhvervs erfaring er vigtigere end sprog, men da vores indikator for sprogvanskeligheder er baseret på en ekstrapolation fra en lille stikprøveundersøgelse, er resultatet usikkert.

1 Baggrund

På foranledning af Social- og Integrationsministeriet har KORA, Det Nationale Institut for Kommuners og Regioners Analyse og Forskning (det tidligere AKF), udarbejdet nærværende analyse af, hvilke faktorer der har betydning for ikke-vestlige indvandrere, deres efterkommere samt etniske danskeres tilknytning til arbejdsmarkedet.

Baggrunden for analysen er, at der fortsat er en væsentligt svagere tilknytning til arbejdsmarkedet blandt indvandrere og efterkommere med ikke-vestlig oprindelse set i forhold til danskere, som fx er dokumenteret af Social og Integrationsministeriet (2011). Konkret viser registeropgørelsen, at 55% af de mandlige ikke-vestlige indvandrere var i arbejde i januar 2010, mens det var tilfældet for 68% af de mandlige efterkommere og 79% af de danske mænd. For kvinderne er andelen henholdsvis 45% for indvandrere, 64% for efterkommere og 75% for danskere. Skævheden afspejles også i andelen, der modtager overførselsindkomster, såsom kontanthjælp eller førtidspension, hvor ikke-vestlige indvandrere også er overrepræsenterede.

Vi kender ikke de præcise årsager til indvandrerens lavere tilknytning til arbejdsmarkedet, men nogle af de individuelle faktorer, vi undersøger i denne rapport, omfatter bl.a. uddannelse og sproglige kompetencer, men også psykisk og fysisk helbred, der kan være afgørende for flygtninge, som kommer til landet med traumatiske oplevelser i bagagen.

KORA har for nylig udgivet en undersøgelse, der konkluderer, at indvandreres medbragte uddannelse udnyttes for dårligt på arbejdsmarkedet (Arendt, Nielsen & Jakobsen 2012). Undersøgelsen er baseret på en spørgeskemaundersøgelse fra 2006 blandt 761 18-45-årige indvandrere fra Tyrkiet, Pakistan og Iran, der er kommet til Danmark som voksne flygtninge eller familiesammenførte før år 2000. Tallene viser, at mindre end halvdelen af indvandrerne helt uden uddannelse har et job, mens det er knap to tredjedele blandt de indvandrere, der har en uddannelse med fra hjemlandet. I modsætning hertil har 90% af de indvandrere, der har taget en dansk videregående uddannelse, og 75% af dem, der har taget en dansk erhvervsfaglig uddannelse, et job.

Utilstrækkelige sprogkundskaber kan være en af årsagerne til, at de personer, der har medbragt en uddannelse fra hjemlandet, ikke i lige så høj grad har succes på arbejdsmarkedet, som de personer, der tager en dansk uddannelse. Her peger undersøgelsens resultater på, at gode dansk kundskaber generelt gør det nemmere at få et job for personer med medbragt uddannelse, men også at det ikke er nok til at give bedre muligheder end for personer helt uden uddannelse.

Med hensyn til det mentale helbred, så viser en opgørelse fra Danmarks Statistik (2011), at flygtninge fra krigshærgede områder på Balkan og i Mellemøsten er overrepræsenterede i førtidspensionsstatistikken sammenlignet med ikke-vestlige indvandrere fra fredeligere områder. Vi ved, at godt halvdelen af alle tilkendelser af førtidspension sker på baggrund af en psykisk lidelse, men for personer med ikke-vestlig herkomst var psykisk sygdom årsagen til knap trefjerdele af nye tilkendelser i 2009. I over halvdelen af disse tilfælde var diagnosen 'posttraumatisk stresssyndrom' (Ankestyrelsen 2010). Diagnosen er bl.a. kendetegnet ved

koncentrationsbesvær og irritabilitet, men også ved at personen i stressende situationer kan reagere med frygt og hjælpeløshed. En international undersøgelse har vist, at prævalensen af post-traumatisk stresssyndrom er 10 gange højere blandt flygtninge end den aldersmæssigt tilsvarende stedlige befolkning (Fazel, Wheeler & Danesh 2005).

Nærværende rapport undersøger, hvilken rolle en række faktorer isoleret set spiller, for indvandrernes tilknytning til arbejdsmarkedet. Vi ser bl.a. på, hvor meget uddannelse betyder under hensynstagen til en række forhold, der ellers kunne tænkes at have betydning. De forhold, der inddrages i denne analyse, omfatter bl.a. køn, alder, civilstand, børn, herkomst, sprogproblemer, opholdsgrundlag, erhvervs erfaring i Danmark samt en række oplysninger om borgerens psykiske og fysiske helbred. Oplysningerne dækker hele befolkningen og stammer fra administrative registre. Vores fokus er på objektive karakteristika ved individet, og vi undersøger derfor ikke, hvordan subjektive faktorer (som eksempelvis motivation) eller overordnede samfundsforhold, såsom lovgivningen, påvirker tilknytningen til arbejdsmarkedet.

Når vores resultater viser, at en ikke-vestlig indvandrер med en erhvervsfaglig eller videregående uddannelse har en 30-50% større chance for at være beskæftiget end en person med kun en grundskoleuddannelse, så betyder det ikke nødvendigvis, at man ville kunne opnå en betydelig beskæftigelsesmæssig *effekt* ved at indføre en politik, der øger uddannelsesniveaue t blandt indvandrere. For at drage slutninger om kausale sammenhænge skal man også tage højde for, at personer, som tager en uddannelse, i udgangspunktet kan være forskellige fra personer, der ikke uddanner sig. Kort sagt, så skal resultaterne i denne rapport derfor tolkes som statistiske sammenhænge, korrelationer, der kan bruges i det videre arbejde med at kortlægge de egentlige årsagssammenhænge bag lav tilknytning til arbejdsmarkedet.

I næste afsnit beskriver vi det datamateriale og den analysemetode, der er grundlaget for rapportens resultater. Derefter præsenterer og diskuterer vi analyseresultaterne. Rapporten efterfølges af en række bilag, der bl.a. viser følsomhedsanalyser.

2 Data og metode

Vores formål er at undersøge, hvilke karakteristika ved individet der forklarer tilknytningen til arbejdsmarkedet for indvandrere og efterkommere med ikke-vestlig herkomst samt personer med dansk oprindelse. I dette kapitel beskrives, dels hvordan populationen i undersøgelsen er udvalgt dels de registre mv., der ligger til grund for de forklarende variable, som vi har med i analysen af tilknytningen til arbejdsmarkedet. Vi diskuterer også konstruktion af de vigtigste variable, der indgår i de regressioner, som vi præsenterer i de følgende afsnit. Kortfattet beskrives også den statistiske metode, vi anvender, og hvordan resultaterne skal fortolkes. I sidste del af afsnittet gennemgås fordelingen af de variable, der er benyttet i analyserne, og vores forventninger om, hvordan de påvirker tilknytningen til arbejdsmarkedet, diskuteres.

Datamaterialet omfatter alle personer bosiddende i Danmark henholdsvis i år 2006 og år 2009. Da vores fokus er tilknytning til arbejdsmarkedet to år senere, udvælger vi personer, der to år senere er i alderen 25-59 år. Normalt medregnes alle 16-64-årige i arbejdsstyrken, men vi ser kun på personer mellem 25 og 59 år, idet erhvervsfrekvensen er højst for denne gruppe.¹ Personer på en permanent overførselsindkomst er ikke inkluderet i undersøgelsen, fordi de har forladt arbejdsmarkedet for bestandig (og vender kun tilbage i ubetydelig grad). Vi har defineret modtagere af 'permanent overførselsindkomst' som førtidspensionister og personer, der modtager fleksjobrelaterede ydelser. Personer under uddannelse eller personer på barsels- eller sygeorlov er inkluderet i analysen, fordi de vil eller kan indtræde i arbejdsstyrken på et senere tidspunkt (selvom de ikke for øjeblikket står til rådighed).

Indvandrere og efterkommere har oprindelse i ikke-vestlige lande, som bl.a. omfatter lande i Mellemøsten, Østeuropa (tidligere østbloklande), Asien, Afrika og Sydamerika. Vi benytter samme definitioner som Danmarks Statistik. Således er en indvandrer defineret ved at være født i udlandet. Ingen af forældrene er både danske statsborgere og født i Danmark. Hvis der ikke findes oplysninger om forældrene, og personen er født i udlandet, opfattes personen som indvandrer. Efterkommere er derimod født i Danmark. Ingen af forældrene er både danske statsborgere og født i Danmark. Findes der ingen forældreoplysninger, og personen er udenlandsk statsborger, opfattes personen som efterkommer. Personer med dansk oprindelse er personer, uanset fødested, der har mindst en forælder, der både er dansk statsborger og født i Danmark. Den udvalgte population for 2009 omfatter godt 2,3 mio. personer, hvoraf 92% er af dansk oprindelse, 7,4% er ikke-vestlige indvandrere, mens 0,6% er efterkommere.

Som det beskrives nedenfor, operationaliserer vi bl.a. et individs tilknytning til arbejdsmarkedet ved at se på lønnen og risikoen for at gå på førtidspension. For at sikres os mod, at de variable, hvormed vi vil forklare lønindkomsten, faktisk er konsekvenser af lønnen snarere end det omvendte – fx kan dårligt helbred være en konsekvens af lav løn lige så vel som det

¹ I 2011 var erhvervsfrekvensen 72-85% for aldersgrupperne mellem 25 og 59 år, mens den for 16-24-årige var 46-64% og 47% for 60-64-årige. Kilde: Statistikbanken.dk.

kan være årsag til lav løn – så bruger vi indkomstoplysninger, der minimum er et år nyere end de forklarende variable.² Havde vi brugt data fra samme år, så ville vi ikke kunne kende forskel på konsekvens og årsag.³ De senest tilgængelige data for lønindkomst (pr. juni 2012) er opdaterede indtil medio 2011.

De forklarende variable, vi benytter i analysen, kommer fra Danmarks Statistiks IDA-register, Lægemedelstyrelsens Lægemedelstatistikregister, Sundhedsstyrelsens Landspatientregister samt Sygesikringsregisteret. Udfaldsvariablene er konstrueret på baggrund af information fra eIndkomstregisteret og Arbejdsmarkedsstyrelsens DREAM-database. Fra IDA trækker vi oplysninger om alder, køn, civilstand, hjemmeboende børn, bopæl, uddannelse, stilling, hovedbeskæftigelse, sektor, erhvervsgrad, årsledighedsgrad, langvarigt sygefravær, indvandringsdato og oprindelsesland.

Ud fra de oplysninger om hver borgers årsledighedsgrad (se beskrivelse i næste kapitel), der ligger i vores datasæt, genererer vi en variabel, der måler, hvor høj den lokale arbejdsløshed er i kommunen, hvor borgeren bor. Vi genererer også en variabel, der måler, hvor stor en andel ikke-vestlige indvandrere udgør af befolkningen i bopælskommunen. For hver person i vores datasæt angiver disse to variable, i hvilket kvintil personens bopælskommune ligger.

Vi er også interesserede i at kontrollere for, om en indvandrer er flygtet til Danmark. Fra Danmarks Statistiks befolkningskontor har vi fået leveret oplysninger om indvandrernes opholdsgrundlag. Disse oplysninger er dannet ved dels at se på faktuelle oplysninger, hvis de eksisterer, og dels ved at se på, hvilket år en person er indvandret, og hvilket land hun kom fra. Variablen er nærmere beskrevet i Danmarks Statistiks publikation ”Indvandrere i Danmark 2008”, som findes online. På den baggrund kan vi opdele indvandrerne i dem, der er flygtet hertil, og restgruppe bestående af alle andre.

Lægemedelstatistikken indeholder oplysninger om salget af receptpligtige lægemidler i Danmark. Lægemidlerne er kategoriseret efter WHO's adaption af det norske ATC-system. Denne anatomisk-terapeutiske klassifikation inddeler kroppen i en række forskellige systemer og klassificerer lægemidlet efter, hvilket system det har virkning på. Vores datasæt indeholder ikke oplysninger om alle ATC-koder, men kun dem der vurderes at være vigtige i en arbejdsmarkedssammenhæng.

I Landspatientregisteret finder man oplysninger om udskrivinger af personer, der har været i behandling på et hospital. Patienterne kan enten have været indlagt, i ambulant behandling eller på skadestuen. Vi benytter ikke oplysninger om sidstnævnte i denne rapport. Som helbredsindikator bruger vi oplysninger om den aktionsdiagnose, der er tilknyttet udskrivingen. Hver person kan altså godt have flere forskellige diagnoser over et år, men hver enkelt udskriving har kun én aktionsdiagnose. Diagnoserne er klassificeret efter det internationale ICD-10 system.

² De forklarende variable er oftest målt pr. november 2009, mens lønnen fx måles mellem januar og juni 2011.

³ Selvom vi har lagget de forklarende variable i forhold til arbejdsmarkedstilknytningen, så kan vi alligevel ikke fortolke resultaterne som årsagssammenhænge. De laggede variable er nødvendige for at sikre os imod omvendt kausalitet, men de er ikke tilstrækkelige til, at sammenhængen er kausal, da der stadig kan være uobserverede sammenhænge, som vi ikke har kontrolleret for.

Oplysninger om kontakter til læger (ydere) under den offentlige sygesikring er klassificeret efter lægens speciale.

De individbaserede lønudbetalinger inklusive ATP-bidrag og personalegoder stammer fra eIndkomstregisteret. Data er rapporteret månedligt og er tilgængelige fra og med januar 2008 til og med juni 2011. Vi benytter kun lønsummen for januar til juni 2011 og fra januar til juni 2008 (begge måneder inklusive). I vores hovedanalyse analyserer vi, hvorvidt summen af de registrerede lønudbetalinger i første halvår af 2011 overstiger et defineret niveau. I en sekundær analyse ser vi på lønnen i perioden januar til juni 2008. Sidstnævnte analyse repræsenterer perioden, før krisen slog i gennem, mens hovedanalysen repræsenterer perioden under krisen.⁴

For at undersøge, hvad der kan være forklaringen på den større tilbøjelighed til, at indvandrerne overgår til førtidspension, bruger vi DREAM-databasen (version 28). Denne database indeholder ugentlige oplysninger for borgere, der modtager offentlige overførsler. For vores 2006-udtræk af 25-59-årige i arbejdsstyrken ser vi således på, i hvilket omfang de overgår til førtidspension (kode 783) i perioden 2007 og indtil primo 2012. Sekundært benyttes DREAM også til at danne en indikator for, om en person i en given uge er selvforsørget eller ej (analysen findes i bilag 4). En person defineres som værende selvforsørgende, hvis hun ikke modtager nogen offentlig overførselsindkomst.⁵

Medborgerskabsundersøgelsen er den sidste datakilde, vi anvender. Det er et survey foretaget blandt et repræsentativt udsnit af befolkningen i 2010-11. Fra denne undersøgelse udtrækker vi godt 2.000 besvarelser angående den ikke-vestlige indvandrerbefolknings selv-vurderede sprogproblemer. Spørgsmålene handler om folks (selvvurderede) evne til at forstå det danske sprog i medierne, i kontakten med kommunen og på arbejdet, men også interviewerens registrering af eventuelle kommunikationsproblemer under samtalen.

Ud fra disse besvarelser har vi konstrueret en indikatorvariabel, som angiver, om borgeren har oplevet sprogproblemer eller ej. Denne variabel linkes via CPR-nummeret med ovennævnte registeroplysninger. Derpå kører vi en regression, hvor vi forklarer sprogvariablen med nøjagtig de samme forklarende variable, som vi bruger i vores øvrige analyser – blot består stikprøven kun af 2.000 personer, som er et subset af vores samlede population. På baggrund af resultaterne kan vi give et kvalificeret bud på, i hvor høj grad hver person i hele indvandrerpopulationen oplever vanskeligheder med det danske sprog, selvom vi altså kun har faktiske sprogoplysninger fra et meget lille udsnit. Vi går her ikke i detaljer med denne model, men henviser den interesserede læser til Bilag 1: Konstruktion af sprogindeks.

Alle analyser i denne rapport benytter binære logistiske regressionsmodeller. Ideen bag en regressionsanalyse er illustreret i Figur 2.1. Figuren viser, at forskellige forhold kan påvirke arbejdsmarkedstilknytningen. Vores undersøgelse omfatter kun et mindre udsnit af alle

⁴ Den sæsonkorrigerede arbejdsløshedsprocent steg fra et niveau på omkring 3,3% i forsommeren 2008 til omkring 6% i forsommeren 2009. Fra efteråret 2009 og indtil omkring april 2010 steg arbejdsløsheden igen fra omkring de 6% til godt 7,5%. Kilde: Eurostat og Google Public Data.

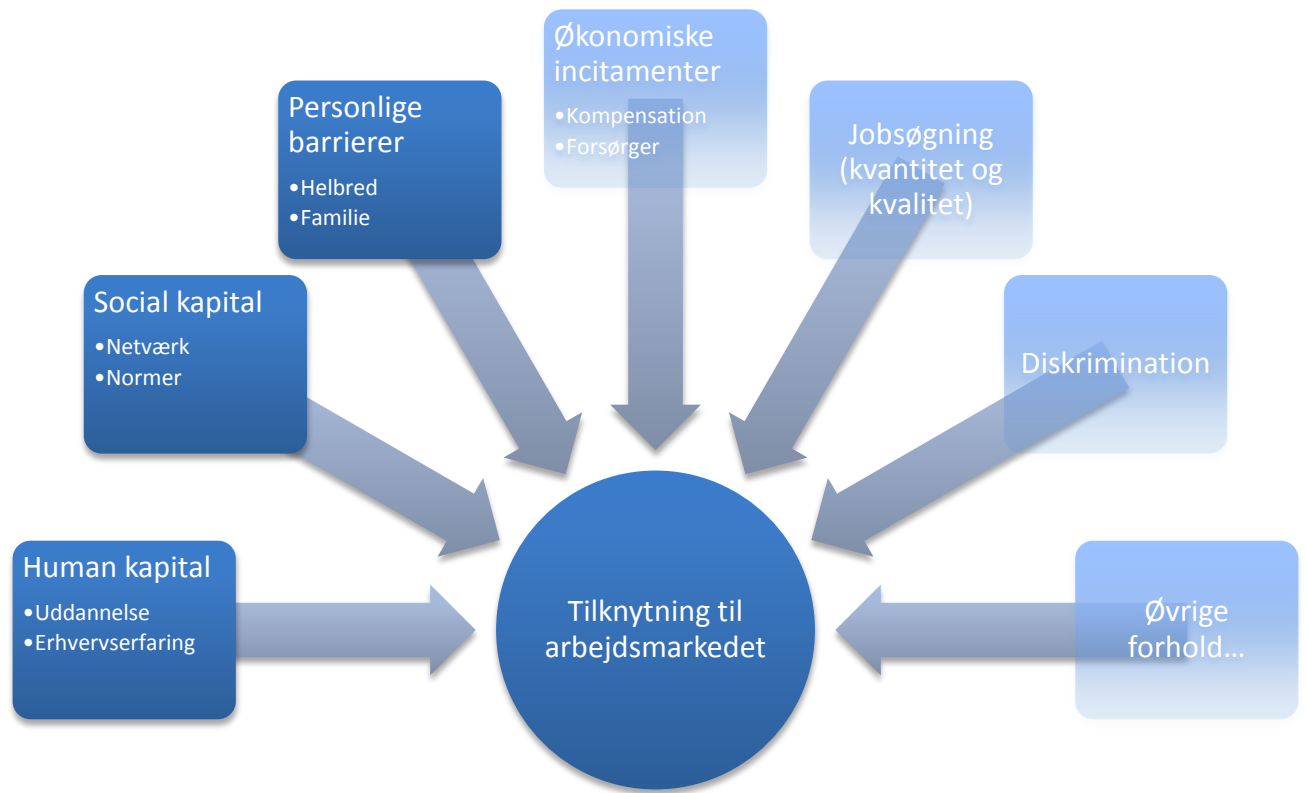
⁵ Det er vigtigt at være opmærksom på, at marginaliserede befolkningsgrupper, der i registeret optræder som selvforsørgende (dvs. uden offentlig overførselsindkomst), ikke nødvendigvis er lønmodtagere; de kan i højere grad være forsørgt af ægtefællen, arbejde sort eller bedrive kriminalitet.

de tænkelige forhold, der kan være bestemmende for, om en person arbejder eller ej. Vores fokus ligger på karakteristika ved individet selv (illustreret med en mørkere farve i figuren) og fx ikke efterspørgselsforhold eller politisk bestemte forhold såsom kompensationsgraden eller den aktive arbejdsmarkedspolitik.

I en binær model er udfaldsvariablen (også kaldet 'y' eller venstresidevariablen) altid lig med nul eller en. Udfaldsvariablen er den variabel, der beskriver personens arbejdsmarkeds-tilknytning. De forklarende variable (ofte kaldet 'x' eller højresidevariablen) kan i princippet både være binære og kontinuerte. Et eksempel på en binær forklarende variabel er køn, som er lig med en vis person og er en mand, og nul hvis det er en kvinde. Alder er et godt eksempel på en variabel, der kan modelleres kontinuert. For at gøre fortolkningen lettere har vi dog valgt, at modellere stort set alle forklarende variable som binære indikatorer. Således inkluderer vores modeller seks indikatorer for femårsaldersgrupper (25-29, 30-34, 40-44, 45-49, 50-54 og 55-59), mens den syvende kategori (35-39 år) tjener som reference. At være reference betyder, at vores resultater viser, hvordan eksempelvis en 50-54-årig person klarer sig på arbejdsmarkedet relativt til en 35-39-årig.

Logistiske regressionsmodeller benyttes, når man skal modellere sandsynligheder. I vores tilfælde er vi interesserede i at vide, hvordan en forklarende variabel påvirker sandsynligheden for, at en person tjener over et vist beløb, ikke har nogen lønindkomst eller tilkendes førtidspension mv. I vores analyser rapporterer vi såkaldte odds-ratioer, da de er nemmest at fortolke. Hvis 55% af alle mandlige indvandrere eksempelvis er i beskæftigelse, mens det er 50% af kvinderne, så vil odds-ratioen for mænd være 1,1. Hvis odds-ratioen er 1,0, så betyder det, at sandsynligheden er den samme i begge grupper. Vi vil ikke gå i detaljer med den statistiske metode, vi benytter, men i stedet henvise den særligt interesserede læser til den engelsksprogede litteratur på området (fx Wooldridge (2009) eller lignende).

Figur 2.1 Illustration af ideen bag en regressionsanalyse



3 Beskrivende statistik

Tabel 3.1 viser fordelingen af de demografiske og socioøkonomiske variable, som vi bruger i analysen. Tabellen viser, at de ikke-vestlige indvandrere og deres efterkommere generelt er yngre end personer med dansk oprindelse.⁶ Eksempelvis er knap 18% af de ikke-vestlige indvandrere i alderen 25-29 år, mens det er tilfældet for 43% af efterkommerne mod kun knap 12% af danskerne. Generelt er flere indvandrere i aldersgrupperne 25-34 år, mens færre er i aldersgrupperne 50-59 år sammenlignet med danskerne. Samtidig er andelen af mænd lidt højere blandt danskere end blandt indvandrere.

Beskæftigelse, alder og køn hænger sammen af mange forskellige grunde. Blandt andet på grund af uddannelse og helbred er det sådan, at helt unge og ældre har lavere erhvervsfrekvens. Vi kontrollerer både for uddannelse og helbred, men forventer stadig, at unge og ældre har lavere arbejdsmarkedstilknnytning, hvis fx disse grupper har større præferencer for fritid frem for arbejde.

For danskernes vedkommende er knap 48% af populationen ikke gift, mens det kun er tilfældet for godt 37% af indvandrerne på trods af, at de er yngre. Personer, der ikke er gift, er enten enlige (dvs. ugifte eller samboende), fraskilte, enker mv. Ægteskabsfordelingen kan være udtryk for andre kulturelle traditioner og værdier, og således kan det også være et mål for et mere konservativt familie- eller kønsrollemønster. Man må også forvente, at det er en selekteret gruppe af personer, som finder en partner at indgå ægteskab med, hvorfor denne selektion vil komme til udtryk i resultaterne. Denne selektion kan ske på baggrund af karakteristika som fx evner og sociale kompetencer, der ellers kan være uobserveret i vores registerdata, men som er potentielt vigtige for arbejdsmarkedstilknnytningen. Eksempelvis for mandlige indvandrere kunne en hypotese også være, at ægteskabet fører til forsørgerforpligtelser, hvorfor vi vil forvente, at ægteskab påvirker beskæftigelsen i positiv retning. Da flertallet af populationen har indgået ægteskab, så er gifte personer referencegruppen i analyserne.

Ligeledes kan fertiliteten udtrykke både selektion og forsørgerforpligtelse. Blandt indvandrerne ser vi en større andel med mange hjemmeboende børn under 18 år. Hvor andelen af indvandrere med tre eller fire og flere hjemmeboende børn er henholdsvis 11,3% og 6,4%, er de tilsvarende tal for danskerne 7,6% og 1,3%. Vi forventer, at indvandrerkvinder med mange børn har en lavere beskæftigelse, fordi deres samlede nytte af passe børnene hjemme overstiger den hypotetiske lønindkomst fraregnet udgifter til børnepasning. På grund af negativ selektion forventer vi, at personer, der ikke har børn, har en lavere beskæftigelse.

⁶ Bemærk, at når vi taler om 'danskere', så mener vi personer med dansk oprindelse. Dette gælder i hele rapporten.

Med hensyn til årsledighedsgrad, måles det som den andel af året (2009), hvor personen har været ledig, relativt til den andel af året, hvor personen har været i arbejdsstyrken.⁷ Vi forventer, at personer med højere årsledighed efterfølgende har en lavere beskæftigelse og løn primært på grund af, at der er tale om en selekteret gruppe, der også de tidligere år har haft høj ledighed, men også fordi markedsværdien af deres kompetencer forringes under ledighedsperioden. Erhvervsgraden måler antal års erhvervserfaring siden 1980. Unge samt personer, der er indvandret for nylig, har lavere erhvervserfaring, men disse forhold kontrolleres der for i regressionerne, hvorfor eventuelle niveauforskelle opfanges af fx indikatoren for, hvor længe en indvandrer har opholdt sig i landet. Fordelingen af erhvervsgraden i data for danskere er nærmest uniform, mens den for indvandrere er meget skæv, sådan at kun få har en mellemhøj eller høj erhvervsgrad.

Sygedagpenge måler, om en person har haft sygefravær ud over arbejdsgiverperioden, som dengang på var to uger. Det betyder, at vi kun måler sygefravær, der er af længere varighed. Flere indvandrere og efterkommere har mere langtidssygefravær end danskere og har dermed fået udbetalt sygedagpenge. Hvis man har haft et langt sygefraværsløb, kan det både skyldes dårligt helbred, men fordi der er tale om en forsikringsordning, kan der også være tale om personer, der har høje præferencer for fritid. I den engelsksprogede forsikringslitteratur er der to centrale begreber: 'adverse selection' og 'moral hazard'. Den moralske hazard opstår, fordi forsikringstageren har et lavere incitament til at være agtpågivende, når han allerede har tegnet en forsikringspolice. Men han har ikke kun laver incitament til at være forsigtig med sit helbred. Han får også et incitament til at snyde sit forsikringsselskab. Da forsikringsselskabet i dette tilfælde er staten, og da forsikringen dækker alle danskere (den er universel), så er der ikke nogen advers selektion. Advers selektion optræder, hvis der er en systematik i, hvilke borgere der tegner en forsikring.

Vi inkluderer to variable, der indikerer, om borgeren bor i en kommune med en høj ledighed eller en stor andel af indvandrere. En høj lokal ledighed kan være et signal om en lav efterspørgsel efter arbejdskraft, mens en høj andel af indvandrere kan være et signal om et rummeligt arbejdsmarked, eller det kan indikere tilstedeværelsen af sociale netværk, som kan bruges i beskæftigelsessammenhæng. Vi kontrollerer samtidig for, om personen bor i København eller på Frederiksberg, i en af kommunerne på den københavnske vestegn, eller i Aarhus, Aalborg og Odense, da koncentrationen af indvandrere i disse kommuner er højere end i de øvrige provinskommuner. Den højere indvandrerkoncentration kan være udtryk for en positiv selv-selektion, sådan at indvandrere, der er mere motiverede til at finde et job, fx bosætter sig i en kommune med et større arbejdsmarked. I analysen af førtidspension skal kommunevariablene snarere tolkes som udtryk for forskellig lokal tilkendelsespraksis.

⁷ Se nærmere beskrivelse af variablen "arledgr" i Danmarks Statistiks onlinedokumentation. (<http://www.dst.dk/da/Statistik/dokumentation/Times>). I regressionsanalysen kan det være et problem, at årsledighedsgraden for personer, der har en lavere arbejdsmarkedstilknytning, er undervurderet. Således har fx 80% af kontanthjælpsmodtagerne og 83% af de øvrige uden for arbejdsstyrken en årsledighedsgrad på nul. Til sammenligning gælder dette 'kun' for 92% af lønmodtagerne på mellem- og grundniveau. Vi ville altså forvente, at årsledighedsgraden var relativt højere for kontanthjælpsmodtagere og øvrige uden for arbejdsstyrken. Men da vi i tillæg kontrollerer for disse stillingskategorier, så skulle niveauforskellene blive opfanget af disse indikatorer.

Bemærk, at kommunens indvandrerandel og variabelen 'Lokal ledighed' ikke, som de øvrige variable, skal tolkes som procent, da de er opgjort i kvintiler. Et 1-tal betyder således, at borgeren bor i en kommune, hvor ledigheden er blandt den laveste femtedel, mens et 5-tal betyder, at ledigheden er i den højeste femtedel.

Før vi taler indvandrere og uddannelsesstatistik, er det nødvendigt at bemærke, at statistikken er ufuldkommen: Der er en høj andel indvandrere – godt 22% – for hvem der ikke findes oplysninger om den højeste fuldførte uddannelse. I analyserne kontrollerer vi for det ved hjælp af en indikatorvariabel, som er lig med en, hvis der ikke foreligger nogen oplysninger om uddannelsesniveaue, og ellers nul.

For både indvandrere og efterkommere er der et klart uddannelsesmæssigt efterslæb relativt til danskerne. Således har 28% af indvandrerne og 35% af efterkommerne højst fuldført en grundskoleuddannelse. Imidlertid så vi ovenfor, at indvandrerne er meget yngre, hvorfor de muligvis med tiden indhenter efterslæbet. For efterkommerne er der imidlertid stort set lige så høj en andel, som har fuldført en lang videregående uddannelse, som hos danskerne – knap 1/10. Vi forventer, at uddannelse både er et signal om produktivitet og evner, men også at uddannelse giver kompetencer, som efterspørges af markedet.

Tabel 3.1 Fordeling af demografiske og socioøkonomiske variable (2009)

	Danskere	Indvandrere	Efter- kommere
	Procent		
Alder: 25-29 år	11,7	17,8	43,0
- 30-34 år	12,7	18,0	35,1
- 35-39 år (B)	15,1	17,2	18,2
- 40-44 år	15,5	16,8	2,1
- 45-49 år	16,8	14,5	0,5
- 50-54 år	14,5	9,8	0,6
- 55-59 år	13,7	5,9	0,5
Køn: Mand	50,9	47,5	50,5
Civilstand: Ikke gift	47,6	37,3	54,3
Ingen hjemmeboende børn	46,1	43,0	51,5
- 1 barn	20,2	18,4	18,5
- 2 børn (B)	24,8	20,9	20,2
- 3 børn	7,6	11,3	7,5
- 4 børn eller flere	1,3	6,4	2,3
Årsledighedsgrad i 2009: Ikke ledig (B)	87,3	80,5	77,4
- Mellem 1 og 19%	6,4	8,7	10,1
- 20% eller derover	6,3	10,8	12,4
Erhvervsgrad fra 1980: Under 33% (B)	35,8	86,0	88,6
- Mellem 33 og 66%	32,3	12,1	10,7
- Over 66%	31,9	1,9	0,7
Har modtaget sygedagpenge i 2009	13,6	15,7	18,5
Bopælskommune: København	12,6	25,3	33,6
- Vestegnskommuner	6,8	14,9	26,3
- Aarhus	5,8	7,6	5,8

	Danskere	Indvandrere	Efter- kommere
	Procent		
- Aalborg	3,8	3,1	0,8
- Odense	3,4	4,4	4,4
- Alle øvrige kommuner (B)	67,6	44,7	29,1
* Indvandrerandel i kom. (kvintil; 1/lav, 5/høj)	3,5	4,2	4,6
* Lokal ledighed (do.)	3,2	3,3	3,4
Højeste fuldførte uddannelse: Grundskole (B)	18,7	28,2	34,6
- Gymnasial uddannelse	6,1	9,9	12,1
- Erhvervsfaglig uddannelse	28,0	18,9	18,2
- Kort videregående uddannelse	17,4	6,4	12,7
- Mellemlang videregående uddannelse	15,5	7,3	8,3
- Lang videregående uddannelse	9,3	6,1	9,6
- Ikke kategoriseret uddannelse	3,7	0,6	0,9
- Manglende uddannelsesoplysninger	0,7	22,3	3,5
Stilling: Lønmodtager på mellem- eller grundniv. (B)	50,9	25,0	41,4
- Topleder eller lønmodtagere på højeste niveau	17,5	5,5	9,6
- Øvrige lønmodtagere	11,8	21,4	11,2
- Selvstændig	5,8	5,9	5,2
- Ledige (uge 48)	3,4	5,8	6,8
- Syge- og barselsdagpenge	1,8	3,1	4,2
- Aktivering og rehabilitering	2,3	5,8	4,4
- Kontanthjælp og integrationsydelse	1,6	9,8	3,7
- Øvrige uden for arbejdsstyrken (1)	2,8	15,1	7,7
- Under uddannelse	2,1	2,6	5,9
Antal personer	2.137.467	172.993	13.820

'**' indikerer, at variablen ikke er i procent. (B) henviser til referencegruppen i de efterfølgende regressionsanalyser.

(1) Se definition af 'pstill'-variablen i Danmarks Statistiks 'Høj kvalitetsdokumentation', som findes online.

Hvis sandsynligheden taler for, at en person, der i dag er lønmodtager, også vil være lønmodtager om to år, så virker det sandsynligt, at en persons stilling i 2009 generelt vil være en stærk prædiktør for hans beskæftigelse eller lønindkomst i 2011. Man kan derfor diskutere, hvorvidt en persons stilling er en relevant information at tage med i denne analyse af arbejdsmarkedstilknytningen, fordi 'stilling' netop beskriver beskæftigelsen eller arbejdsmarkedstilknytningen, hvilket samtidig er det, vi forsøger at forklare. På en måde kan man sige, at vi, ved at inddrage stillingsoplysninger, risikerer at forklare en persons arbejdsmarkedstilknytning med hans arbejdsmarkedstilknytning. Spørgsmålet er, om eksempelvis nuværende ledighed forringer et individs fremtidige beskæftigelsespotentiale igennem strukturelle sammenhænge, eller det blot er en ren statistisk sammenhæng, som skyldes uobserverede forhold, der påvirker ledighedsrisikoen, men ikke påvirkes af ledighed i sig selv (Heckman & Borjas 1980). Eksempler på uobserverbare faktorer kan være individets motivation og evner. Begge er korreleret over tid og indfanges ikke nødvendigvis af uddannelse eller nogle af de øvrige kontrolvariable i vores undersøgelse.

For ikke at komplicere den statistiske metode, kan vi enten inkludere stillingsoplysningerne, opdele populationen og analysere fx to grupper af personer, der er henholdsvis tæt på arbejdsmarkedet og lidt længere væk, eller også kan vi ignorere oplysningerne om den nuværende arbejdsmarkedsstatus. Vi afprøver alle tre løsninger, og det viser sig at gøre en forskel for specielt sprogvariablen, om arbejdsmarkedsoplysningerne inkluderes eller ej (se Bilag 3: Alternative modelspecifikationer).

Målt på stillingsstrukturen blandt disse tre grupper, så tegner der sig et billede, der er i overensstemmelse med oplægget bag denne rapport: Andelen af lønmodtagere er lavere blandt indvandrere og deres efterkommere, hvilket modsvares af en højere andel, der modtager overførselsindkomst. Specifikt er 80% af danskerne lønmodtagere, mens kun 52% af de ikke-vestlige indvandrere og 62% af efterkommerne er det. Andelen af selvstændige er stort set ens i de tre grupper, hvilket specielt er bemærkelsesværdigt, når man tager efterkommerne unge alder i betragtning. Når vi ser på ledighed, sygdom, barsel, aktivering og kontanthjælp under et, så tilhører knap 25% af indvandrerne og 19% af efterkommerne denne kategori, mens det er 9% af danskerne. I tillæg befinder 15% af indvandrerne sig i kategorien 'Øvrige uden for arbejdsstyrken', der bl.a. består af personer, som midlertidigt er ude af arbejdsstyrken, fx selvforsørgede. Kategorien dækker primært over personer, som år for år bevæger sig rundt mellem de øvrige kategorier, og som har en svag arbejdsmarkedstilknytning.

I rapportens bilag 4 undersøger vi, hvorvidt demografiske, uddannelses- og helbredsmæssige forskelle kan forklare den forskel i tilknytningen til arbejdsmarkedet mellem danskere og indvandrere, som de simple sammenligninger viser. Det viser sig slet ikke at være tilfældet. Selvom man sammensætter en gruppe af sammenlignelige danskere på de nævnte karakteristika, så mindskes forskellen i løn, selvforsørgelse og overgangen til førtidspension mellem danskere og indvandrere kun ganske lidt (se Bilag 4: Løn, selvforsørgelse og førtidspension blandt ikke-vestlige indvandrere og sammenlignelige danskere). Sammenfattende viser bilagsundersøgelsen, at det ikke er forskelle i demografiske og helbredsmæssige forhold, der er hovedårsagen til forskellen på, hvordan indvandrere med ikke-vestlig herkomst og danskere klarer sig på arbejdsmarkedet. Forskellige demografiske og helbredsmæssige karakteristika er en del af forklaringen, men kun en lille del. Resultaterne indikerer derimod, at sprogmæssige barrierer udgør en større og vigtigere forhindring for en gruppe af indvandrere, om end vi ikke kan udelukke, at de sprogmæssige indikationer blot er udtryk for en underliggende uobserveret sammenhæng.

Også branche- og beskæftigelsesfordelingen i Tabel 3.2 afspejler indvandrerne lavere tilknytning til arbejdsmarkedet. Det ses i og med, at 48% af indvandrerne ikke kan tilknyttes en branche, og 44% ikke har en kategoriseret hovedbeskæftigelse. De dominerende brancher for beskæftigede indvandrere er den offentlige sektor samt handel og transport, og hovedbeskæftigelsen er ofte detailsalg, service- og omsorgsarbejde samt transport. Danskere og efterkommere skiller sig især ud ved at have beskæftigelse, der kræver viden og specialisering på et højere niveau. Indvandrerne er også karakteriseret ved i lavere grad at arbejde i bygge- og anlægssektoren og med industriel fremstilling.

Tabel 3.2 Branche og hovedbeskæftigelse

	Danskere	Indvandrere	Efter- kommere
	Procent		
Sektor: Ingen branchekode (B)	19,9	48,2	38,0
- Landbrug	0,8	1,9	0,1
- Industri	11,3	7,2	4,8
- Byggeri og anlæg	4,7	0,8	1,4
- Handel og transport	16,0	13,2	17,9
- Kommunikation og it	3,3	1,3	3,1
- Finans og forsikring	3,1	0,7	2,6
- Ejendomshandel og udlejning	1,0	0,4	0,5
- Erhvervsservice (fx rengøring)	7,4	9,4	7,9
- Offentlig sektor	29,8	15,4	21,9
- Kultur- og fritidsaktiviteter	2,8	1,4	1,9
Hovedbeskæftigelse: Ikke oplyst (B)	15,9	43,8	27,7
- Militært arbejde	0,8	0,1	0,3
- Ledelse på øverste niveau	3,5	0,4	0,9
- Højt specialiseret arbejde	14,7	5,7	9,3
- Arbejde, der forudsætter viden på mellemniveau	21,5	6,5	15,1
- Kontorarbejde	8,5	3,5	10,9
- Detailsalg, service- og omsorgsarbejde	13,5	13,5	19,4
- Arbejde i de primære erhverv	1,0	0,1	0,1
- Håndværkspræget arbejde	8,5	3,1	3,7
- Maskin-, transport- og anlægssarbejde	5,4	7,3	4,1
- Anden beskæftigelse	6,7	15,9	8,5
Antal personer	2.137.467	172.993	13.820

(B) henviser til referencegruppen.

Tabel 3.3 viser fordelingen af de fire indvandrespecifikke variable, som ikke er defineret for populationen af danskere. Den største gruppe indvandrere – knap 25% – har boet i Danmark i højst fire år, mens 11% har været her i 20 år eller længere. 12% har en ukendt indrejsedato, hvilket oftest betyder, at de har været her længere end siden midt i 1980'erne. Den største gruppe af ikke-vestlige indvandrere er tyrkere, som udgør knap 14% af indvandrerpopulationen og 38% af efterkommerne. Når vi samtidig kontrollerer for bl.a. uddannelse, erhvervs erfaring, opholdsgrundlag samt sproglige kompetencer, så kan oprindelseslandet være en indikator for den kultur, en indvandrer medbringer i sin bagage og dermed også et mål for det modsætningsforhold, der eventuelt kan opstå mellem indvandrer og arbejdsmarkedets forventninger, fx i form af diskrimination eller kønsrollemønstre (Dahl & Jakobsen 2005). Blandt unge indvandrere synes denne hypotese dog ikke at finde støtte i fx Goul Andersen (2008).

Personer, som enten er flygtninge eller statistisk vurderes at være flygtet hertil, udgør 36% af indvandrerpopulationen. Under hensyntagen til et måske lavere uddannelsesniveau

og et eventuelt dårligere mentalt helbred, så forventer vi, at flygtninge kan være længere tid om at blive integreret på arbejdsmarkedet.

Fordelingen af de fire sprogvariable er uniform pr. konstruktion. Variablene indikerer sandsynligheden for at have lav, mellemlav, mellemhøj eller høj risiko for at opleve sprogvan- skeligheder. I Bilag 1: Konstruktion af sprogindeks beskriver vi, hvordan dette indeks er kon- strueret. Vi forventer, at bedre sprogkompetencer påvirker beskæftigelsen positivt.

Tabel 3.3 Indrejse, oprindelse, sprog og opholdsgrundlag (2009)

	Danskere	Indvandrere	Efterkommere
	Procent		
Indrejsedato: Ukendt		12,0	
- Højst 4 år siden indrejse		24,7	
- 5-9 år siden indrejse (B)		17,9	
- 10-14 år siden indrejse		21,3	
- 15-19 år siden indrejse		13,2	
- Minimum 20 år siden indrejse		10,9	
Oprindelsesland: Tyrkiet (B)		13,6	38,0
- Iran		4,7	0,9
- Irak		6,5	0,3
- Vietnam		3,8	3,3
- Libanon		4,5	0,9
- Pakistan		4,4	20,5
- Bosnien		5,5	0,2
- Jugoslavien		4,1	11,8
- Makedonien		1,0	1,1
- Kroatien		0,2	0,5
- Serbien-Montenegro		0,3	0,0
- Thailand		3,9	0,5
- Somalia		3,6	0,2
- Sri Lanka		2,9	0,4
- Filippinerne		3,9	1,2
- Marokko		2,1	6,7
- Afghanistan		2,8	0,2
- Kina		3,7	0,8
- Ukraine		2,7	0,1
- Rumænien		2,5	0,2
- Indien		2,2	2,6
- Rusland		1,9	0,1
- Brasilien		1,0	0,2
- Bulgarien		1,0	0,1
- Syrien		0,8	0,5
- Ghana		0,7	0,1
- Nepal		0,6	0,1
- Øvrige ikke-vestlige lande		15,0	8,4
Opholdsgrundlag: Flygtning eller vurderet flygtning		35,8	
- Ikke flygtning (B)		64,2	
Estimeret sprogproblemer: Lav risiko (<25%)		25,0	
- Lav-mellem risiko for sprogproblemer (25 til 49%) (B)		25,0	
- Mellem-høj risiko for sprogproblemer (50 til 74%)		25,0	
- Høj risiko for sprogproblemer (>75%)		25,0	
Antal personer	2.137.467	172.993	13.820

(B) henviser til referencegruppen.

Tabel 3.4 viser fordelingen af de tre inkluderede helbredsindikatorer. Den første blok omfatter receptpligtige lægemidler købt i perioden 2008 til 2009.⁸ Hospitalsdiagnoser⁹ vedrører samme periode, mens oplysningerne om kontakter til læger i det primære sundhedssystem kun vedrører 2009. Det er værd at bemærke, at vores indikatorer ikke viser, hvor mange recepter eller hospitalsudskrivelser en person har haft, men kun om hun har købt medicin eller været indlagt eller ej. Det gør, at det billede, vi tegner, er unuanceret i forhold til, hvor dårligt personens helbred faktisk er. Vi har imidlertid valgt den enkelte strategi, fordi det ikke nødvendigvis er ensbetydende med, at en psykisk sygdom eksempelvis er værre, blot fordi personen indløser flere recepter. Det kan blot være udtryk for, at personer reagerer forskelligt på medicinen. Derudover er de kategorier, vi arbejder med, meget brede, hvorfor vi mener, at det vil give mindre mening at se på antallet af receptindløsninger eller -udskrivelser.

Tallene viser, at godt 27% af danskerne købte antiinflammatoriske (M01) præparater på recept (fx gigtpræparater, men også mildere smertelindrende produkter), mens 12% købte smertestillende præparater (N02) med virkning på nervesystemet (fx hovedpinepiller og stærke opiumspræparater mod migræne). Knap 14% købte medicin i gruppen af psykofarmaka (N05 og N06), der omfatter medicin mod depression, angstdæmpende midler, sovemedicin og antipsykotiske lægemidler. For indvandrere er tendensen den samme, mens andelen er lidt lavere for de gennemsnitligt yngre efterkommere.

De mest typiske hospitalskontakter sker i forbindelse med almindelig undersøgelse eller svangerskab og reproduktion, men også kontakter for knoglesygdomme er meget almindelige både blandt danskere og indvandrere. Indlæggelser i forbindelse med eksterne årsager, fx bilulykker, er almindelige hos danskere og efterkommere, men lidt mindre for indvandrere. Indvandrere er oftere i kontakt med hospitalet på grund af svangerskab, men som vi så ovenfor, så skyldes det indvandrerkvindernes højere fertilitet.

For kontakter til alment praktiserende læger og speciallæger viser fordelingen, at danskerne i gennemsnit går til læge seks gange om året, mens indvandrere og efterkommere gør det lidt oftere. Generelt ved vi, at der er en social skævhed i kontakter til ydere med en højere grad af egenbetaling, hvilket her bl.a. kommer til udtryk ved, at 65% af danskerne besøger en tandlæge mod kun 26% af indvandrerne. Godt 1% af danskerne har besøgt en psykiater, mens andelen er dobbelt så stor blandt indvandrerne.

⁸ Vi henviser til WHO for nærmere information om, hvilke præparater der er inkluderet i kategorierne.
Se http://www.whocc.no/atc_ddd_index/

⁹ Vi henviser til Sundhedsstyrelsens SKS-browser for yderligere oplysninger om diagnosesystemet. Sundhedsstyrelsens SKS-browser findes på <http://medinfo.dk/sks/>

Tabel 3.4 Lægemidler, diagnoser og sygesikringsydelser (2009)

	Danskere	Indvandrere	Efter- kommere
	Procent		
Køb af receptmedicin: Antacid (A02)	8,4	16,0	10,8
- Insulin (A10)	2,0	3,0	1,6
- Antithrombose (B01)	2,7	2,1	0,5
- Hjerteterapi (C01)	0,8	0,7	0,2
- Diuretica (C03)	4,3	2,4	0,9
- Betablokkere (C07)	3,6	3,0	1,3
- Calciumantagonister (C08)	3,0	1,9	0,4
- Renin-angiotensin syst. (C09)	6,0	3,6	0,8
- Lipid-modif. præp. (C10)	4,7	4,0	0,8
- Antireumatica og antiinflam. (M01)	27,2	27,8	22,4
- Analgetica (N02)	12,3	13,4	9,1
- Pyskoleptica og psykoanaleptica (N05 og N06)	13,6	14,9	10,7
- Anti-afhængighed (N07B)	1,0	0,3	0,3
- KOL-præparater (R03)	7,5	6,4	8,0
- Antihistaminer (R06)	7,2	8,2	8,5
Aktionsdiagnose fra hospital: Infektioner (A, B)	1,3	2,4	2,4
- Svulster (C00-D48)	3,2	2,3	1,4
- Blodsygdomme mv. (D50-D89)	0,4	0,6	0,5
- Ernæringssygdomme mv. (E)	2,9	3,5	3,0
- Psykiske lidelser (F)	0,9	0,9	0,9
- Nervesygdomme (G)	2,5	2,3	2,0
- Øre- og øjensygdomme (H)	2,7	2,9	2,2
- Kredsløbssygdomme (I)	3,7	3,1	1,5
- Åndedrætssygdomme (J)	2,2	2,8	3,5
- Fordøjelsessygdomme (K)	5,2	5,1	4,2
- Hudsygdomme (L)	1,8	2,0	3,1
- Knoglesygdomme mv. (M)	9,0	9,0	6,1
- Urinvejssygdomme mv. (N)	6,4	6,5	6,5
- Svangerskabs-, misdannelser mv. (O, P, Q)	8,7	11,7	17,3
- Div. symptomer (mavesmerter, besvimelse mv.) (R)	5,5	7,6	6,2
- Ydre påvirkning (ulykke mv.) (S, T, X, Y)	7,4	5,3	7,3
- Undersøgelse (Z00-Z13)	14,2	13,2	11,7
- Kontakt i forbindelse med reproduktion (Z30-Z39)	8,6	10,7	16,7
- Øvrige forhold (Z20-Z29, Z40-Z99)	5,1	4,3	4,6
Sygesikrings: *Alm. læge (antal kontakter)	6,0	6,5	7,0
- Tandlæge	64,6	26,2	34,5
- Fysioterapi	8,5	6,2	6,2
- Kiropraktik	8,6	2,1	2,9
- Psykologhjælp	4,9	4,1	5,5
- Psykiatri	1,1	2,3	1,5
- Hud- og kønssygdomme	6,1	5,7	9,8
- Øjenlæge	5,2	5,8	4,7
- Ørelæge	5,6	7,3	7,3
- Gynækologi	4,8	5,5	6,2
- Reumatologi	1,1	2,7	1,7
- Intern medicin	0,8	0,9	1,1
- Kirurgi (plastik, ortopædisk mv.)	3,1	4,1	4,0
Antal personer	2.137.467	172.993	13.820

'*' indikerer, at variablen ikke er i procent.

De i Tabel 3.4 nævnte helbredsindikatorer kan i høj grad være korreleret. Samtidig er de – på grund af antallet af indikatorer – svære at fortolke. Derfor har vi konstrueret et sæt alternative indikatorer for tre 'store' sygdomsgrupper. Grupperne er

- 1 psykiske sygdomme,
- 2 knogle-, muskel- og skeletsygdomme samt
- 3 hjerte-kar-sygdomme.

De nævnte sygdomme er nogle af hovedårsagerne, når man ser på, hvorfor borgere tilkendes førtidspension. Vi definerer psykisk sygdom ved, at man enten besøger en psykolog eller en psykiater, har en psykiatrisk diagnose eller indløser tre eller flere recepter for medicin mod psykiske sygdomme (depression, angst mv.). Denne inddeling skal ses som et forsøg på at skabe en samlet gruppe af personer med en eller anden form for psykisk sygdom. Vi anerkender, at løsningen er upræcis og potentielt kan omfatte personer med et meget varieret sygdomsbillede, men vælger alligevel løsningen i et forsøg på at give et enklere og mere samlet billede end de enkeltstående oplysninger om diagnoser, medicinkøb og lægebesøg.

De to øvrige sygdomsgrupper er defineret ved hjælp af samme metode som psykiske sygdomme. Knogle-, muskel- og skeletsygdomme defineres ved tre eller flere køb af receptpligtig antireumatika eller en diagnose for en knoglesygdom. Hjerte-kar-sygdom er defineret ved fire eller flere køb af medicin, der virker på det kardiovaskulære system eller på en diagnose i det tilsvarende spektrum. Antallet af receptindløsninger varierer over de forskellige sygdomsgrupper, fordi den underliggende fordeling af receptindløsninger er meget forskellig. Med de her anvendte definitioner forsøger vi at fange personer, som har et konstant forbrug, og altså ikke personer, der indløser en recept en enkelt gang eller to.

Defineret sådan er der henholdsvis 11,8%, 13,7% og 11,4% af danskerne, der lider af psykiske sygdomme, knogle-, muskel- og skeletsygdomme samt hjerte-kar-sygdomme. For indvandrere er andelen 11,5%, 13,3% og 7,3%, mens det for efterkommere er henholdsvis 10,1%, 8,6% og 2,4%.

Efter at have gennemgået alle de faktorer, som vi har tænkt os at inddrage i analysen i kapitel 4, så står det klart, at indvanderne adskiller sig fra danskerne på en række områder. I gennemsnit er indvanderne yngre, de er oftere gift, de har flere børn, de har oftere kun en grundskoleuddannelse, og de har lavere tilknytning til arbejdsmarkedet. På nogle områder er deres helbred lidt bedre, mens det på andre områder er lidt værre; overordnet mener vi ikke, at forskellene er markante. De analyser, vi præsenterer i næste afsnit, giver imidlertid et langt mere interessant billede end de simple gennemsnit, vi har set ovenfor, fordi resultaterne siger noget om, hvordan de observerede forskelle i en variabel, som fx helbred, påvirker arbejdsmarkedstilknytningen, at vi har kontrolleret for andre vigtige variable, som fx alder.

4 Analyseresultater

I dette afsnit præsenterer vi resultaterne fra vores analyser. Vi præsenterer resultater fra flere forskellige modeller. Overordnet undersøger modellerne, hvilke faktorer der påvirker sandsynligheden for

- at opretholde en lønindkomst over niveauet for en ufaglært,
- ikke at have nogen lønindkomst og
- at overgå til førtidspension.

Vores undersøgelser er baseret på hele sættet af forklarende variable, som vi gennemgik ovenfor i kapitel 3, men indledningsvis præsenterer vi nogle simple modeller, som kun kontrollerer for ganske få karakteristika. I bilagene undersøger vi resultaternes sensitivitet over for korrelation mellem de forklarende variable ved at præsentere flere alternative modeller, som er specificeret med forskellige forklarende variable.

Konkret svarer analyserne på spørgsmålet om, hvordan de forklarende variable varierer med 1) evnen til at opretholde en lønindkomst over 125.000 kr. i første halvår 2011, 2) risikoen for at have en lønindkomst på 0 kr. i første halvår 2011. Begge disse analyser udføres *både* for 2011, som repræsenterer perioden under den økonomiske krise, *og* for 2008, som refererer til perioden før lavkonjunktoren. Vi præsenterer resultater for to tidsperioder både for at undersøge resultaternes stabilitet over tid, men også for at vurdere, om den økonomiske krise kan have påvirket adfærden. Endelig undersøger vi, hvordan de forklarende variable påvirker 3) risikoen for at overgå til førtidspension i perioden 2007 til medio 2012. I de sidste analyser undersøges overgangen til førtidspension for personer, der i udgangspunktet ikke er på førtidspension. For at opnå et tilstrækkeligt datagrundlag, da det trods alt er relativt få personer, som overgår til førtidspension, har vi valgt at undersøge en 5 ½-årsperiode.

Alle modeller er estimeret som logistiske regressioner, og koefficienterne er gengivet som odds-ratioer. En odds-ratio på 1 betyder, at den pågældende faktor ikke ændrer gruppens sandsynlighed set i forhold til referencegruppen, hvorimod en odds-ratio over en indikerer en højere sandsynlighed og vice-versa for en odds-ratio under en. Referencegruppen er i alle modeller gifte, 35-39-årige kvinder, der bor i provinsen med to børn. Kvinderne arbejder som ufaglærte lønmodtagere og har en god arbejdsmarkedstilknytning (høj erhvervsgrad og lav ledighedsgrad). For indvandrere har kvinderne oprindelse i Tyrkiet, og de har boet i Danmark i 5-9 år.

Nedenfor gennemgår vi først resultaterne fra modellerne for henholdsvis en løn over niveauet for en ufaglært arbejder og ingen lønindkomst i første halvår 2011. Dernæst præsenteres modellen for overgang til førtidspension. Til sidst ser vi på resultaterne for 2008-modellerne. Vi kommenterer løbende på resultaternes følsomhed i forhold til de analyser, der præsenteres i bilagene til rapporten.

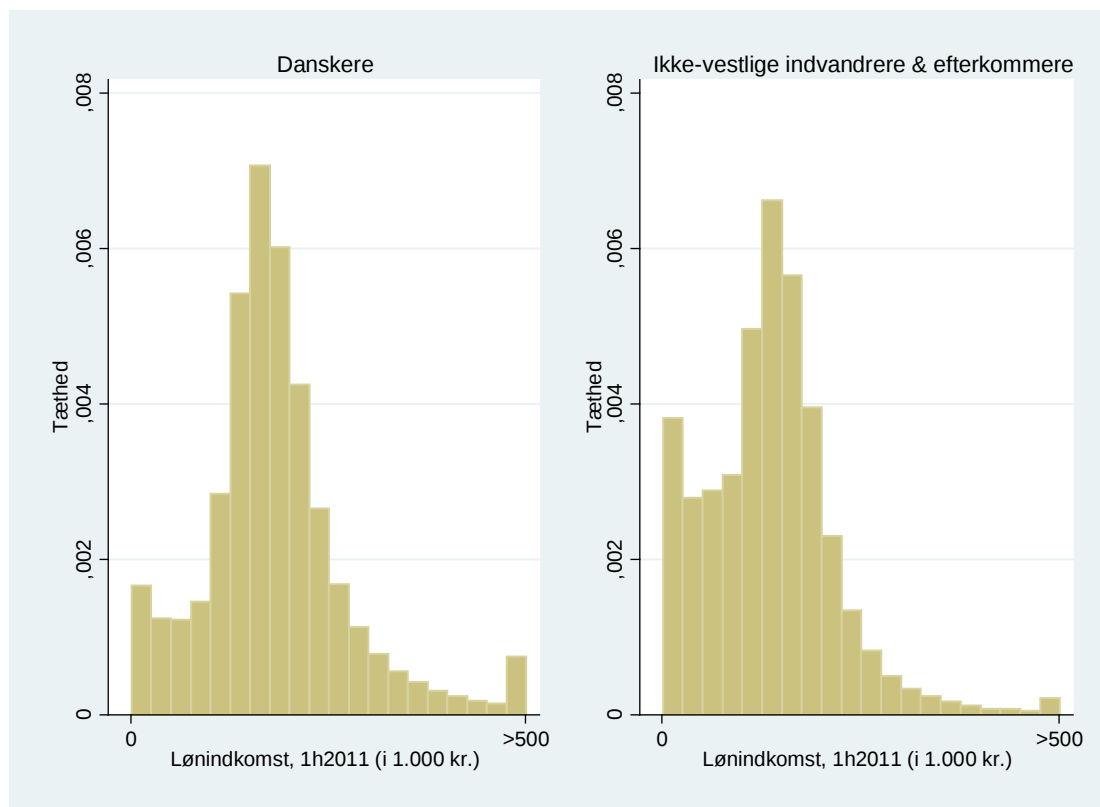
4.1 Modeller for lønindkomst

Dette afsnit viser resultater fra logistiske regressioner af forskellige lønindkomster på de forklarende variable. Konkret er lønindkomst operationaliseret på to måder på baggrund af *summen* af de månedlige lønindkomster for perioden januar til juni 2011 (begge måneder inklusive).

- 1 Den første indikatorvariabel er lig med en, hvis lønnen overstiger 125.000 kr., og ellers nul. Det beløb svarer til en median bruttolønsum for ufaglærte medarbejdere i en børnehave eller vuggestue i vores datasæt. Dette valg kan virke tilfældigt, men vi mener, det repræsenterer et empirisk retvisende niveau for ufaglærte i fast arbejde. For indvandrerne er det kun knap 34%, som opretholder en løn herover, mens det for efterkommere er 42% mod 69% hos danskerne.
- 2 Den anden indikatorvariabel er lig med en, hvis lønnen er lig med nul kroner. Hvor førstnævnte sætter fokus på, hvad der forklarer, om en person faktisk har en lønindkomst, ser vi her på, hvordan samme variabel forklarer, om en person ikke har nogen lønindkomst.

Fordelingen af positive (dvs. større end nul) bruttolønninger i første halvår 2011 fremgår af Figur 4.1, hvor vi ser, at ikke-vestlige indvandrere og efterkommere har en mere skæv lønfordeling end danskerne. Danskernes lønfordeling er tilnærmelsesvis klokkeformet, hvorimod der er større variation blandt indvandrerne.

Figur 4.1 Fordeling af positive lønindkomster blandt danskere og indvandrere



Vi udfører analyserne separat for tre grupper: danskere, ikke-vestlige indvandrere og efterkommere. I tillæg har vi valgt at opdele indvandrergruppen i mænd og kvinder. Det gør vi, fordi det er specielt interessant at undersøge, om forskellige karakteristika påvirker arbejdsmarkedstilknytningen anderledes for indvandrerkvinder end for indvandrermand. Hos indvandrerkvinderne har knap 29% en løn på 125.000 kr. eller mere, mens tallet er 10 procentpoint højere hos indvandrermandene. Hos danskerne er forskellen langt mindre – 67% af kvinderne mod 71% af mændene tjener over 125.000 kr. – hvorfor vi ikke estimerer et kønsopdelt resultat for denne gruppe.

I Tabel 4.1 viser vi resultaterne fra en logistisk regression på vores to lønudfaldsvariable. Disse resultater er baseret på enklere modeller end de resultater, vi efterfølgende præsenterer, idet vi kun kontrollerer for faktorer, som vi vurderer, er de mest grundlæggende karakteristika, som vores datasæt indeholder. Resultaterne er derfor et godt udgangspunkt i og med, at de både giver os mulighed for at se, hvordan personlige karakteristika ændrer forklaringskraft, når der senere inkluderes flere (indbyrdes korrelerede) forklarende variable, men de giver os også lejlighed til at illustrere nogle af de metodiske udfordringer, som vi løser ved at inkludere flere forklaringsfaktorer. De modeller, vi først præsenterer nedenfor, kontrollerer således kun for alder, køn, civilstand, børn, uddannelse og helbred. For indvandrerne kontrollerer vi også for deres selvvaluerede problemer med det danske sprog, samt om de er flygtninge.¹⁰

Tabellens venstre panel viser odds-ratioer for regressionen af de forklarende variable på indikatoren for, om en person har en lønindkomst på 125.000 kr. eller mere. Panelet i tabellens højre side viser, hvad der forklarer sandsynligheden for ikke at have nogen lønindkomst. Vi diskuterer først resultaterne for en lønindkomst på 125.000 kr. eller mere, og derefter resultaterne for nul-lønindkomst.

Resultaterne viser, at unge under 30 år har lavere sandsynlighed for at have en løn på 125.000 kr. eller mere end 35-39-årige. Eksempelvis viser tabellens første række, at odds-ratioen er 0,448 gange lavere for danskere i alderen 25-29 år end danskere i 35-39-årsalderen. Det svarer til, at sandsynligheden for de helt unge er godt 55% (dvs. $1 - 0,448$) lavere. Mandlige indvandrere i alderen 30-34 år har en sandsynlighed tæt på referencegruppen af 35-39-årige mænd, idet odds-ratioen kun marginalt er forskellig fra en.¹¹ Når vi ser på de lidt ældre befolkningsgrupper, stiger sandsynligheden gradvis for at tjene over 125.000 kr. Dette gælder både for indvandrere og danskere. Denne tendens er ikke overraskende, da folk får længere erhvervs erfaring med alderen, og resultatet skal derfor snarere ses som en underliggende korrelation mellem alder og erhvervs erfaring. Resultaterne for alder for efterkommere viser samme mønster, men er mere statistisk usikre. Vi bemærker også, at mænd har en højere sandsynlighed for at tjene over 125.000 kr., mens personer, der ikke er gift, og perso-

¹⁰ I tillæg til de variable, der er vist i tabellen, så kontrollerer vi for, hvor længe indvandrerne har boet i Danmark, samt om personen bor i en stor kommune. Resultaterne for disse variable er ikke vist i Tabel 4.1 af pladshensyn, mens de er inkluderet i de efterfølgende tabeller.

¹¹ Graden af sikkerhed angives med tegnene '+', '*' eller '**', hvor intet tegn er usikkert, mens '**' er meget sikkert. Det laveste statistiske niveau for sikkerhed, som vi er villige til at acceptere, er '+', som betyder, at odds-ratioen er forskellig fra en på et såkaldt 10%'s niveau.

ner med tre eller flere hjemmeboende børn har en lavere sandsynlighed for at tjene over 125.000 kr. sammenlignet med henholdsvis personer, der er gift, og personer med kun to hjemmeboende børn. Kvinder uden hjemmeboende børn har en 14% højere sandsynlighed end kvinder med netop to hjemmeboende børn.

Når vi ser på uddannelse, har danskere, der højst har fuldført en gymnasial uddannelse, en 10% højere sandsynlighed for at tjene over 125.000 kr. end en dansker, der kun har en grundskoleuddannelse. Derimod viser resultatet for indvandrere, at en person med en gymnasial uddannelse har en sandsynlighed for at tjene over 125.000 kr., der er 21% (dvs. $(1-0,787)*100$) lavere end en indvandrer, der kun har en grundskoleuddannelse. Mandlige indvandrere med en erhvervsfaglig uddannelse har ikke en statistisk signifikant anderledes sandsynlighed end indvandrede mænd med kun en grundskoleuddannelse, mens indvandrede kvinder derimod klarer sig bedre med en erhvervsfaglig uddannelse. For danskere har personer med en erhvervsfaglig uddannelse næsten en dobbelt så stor chance for at tjene over 125.000 kr., mens efterkommere har en godt tre gange højere chance. Når vi ser på de mellemlange og lange videregående uddannelser, så har danskere med disse uddannelser mere end dobbelt så høj sandsynlighed for at tjene over 125.000 kr., mens en mellemlang eller lang videregående uddannelse kun øger sandsynligheden med 20-40% for indvandrerne, afhængigt af deres køn. For efterkommerne betyder det mere for sandsynligheden, end vi så det gjorde for danskerne. Endelig viser resultaterne for indvandrerne samlet set, at den store gruppe med manglende uddannelsesoplysninger ikke statistisk set klarer sig meget anderledes end gruppen med kun en grundskoleuddannelse.

Indvandrere, der er flygtet til landet, har en ca. 24% lavere sandsynlighed for at tjene 125.000 kr. eller mere end indvandrere, der ikke er flygtet hertil, mens indvandrere, der ikke har nogen sprogproblemer, har en knap tre gange højere chance for at tjene over 125.000 kr. sammenlignet med indvandrere, der har små sprogproblemer. Personer med mellemstore eller store sprogproblemer har en halvt så stor eller endnu mindre chance for at tjene mindst 125.000 kr. sammenlignet med indvandrere, der kun har små sprogproblemer. Vi har tidligere beskrevet, at sprogvariablen er udtryk for en persons egen opfattelse af sine problemer med sproget, bl.a. i arbejdssammenhæng. Vi ved ikke, hvordan indvandrerne tolker spørgsmålet, når de bliver spurgt, i hvilken grad de oplever sproglige problemer i arbejdssammenhæng. Hvis de ikke kan finde et job, og de selv oplever, at det skyldes sproglig utilstrækkelighed, så vil arbejdsløse indvandrere have større sandsynlighed for at rapportere om sprogproblemer. I de følgende analyser ser vi nærmere på, hvordan sprogvariablen ændrer betydning, efterhånden som vi kontrollerer for flere individuelle karakteristika.

Helbred forringer chancen for at tjene 125.000 kr. eller mere både for danskere, indvandrere og efterkommere. Værst ser det ud for psykisk sygdom, hvor det sænker sandsynligheden med 49-58% i gennemsnit. Stort set sænker psykisk sygdom chancen for at have en løn over 125.000 kr. med lige så meget, som det gør at have mellemstore sprogproblemer eller at være i aldersgruppen 25-29 år. Hvis en dansker lider af en knogle-, muskel- eller skeletsygdom, så har personen 23% lavere chance for at tjene 125.000 kr. eller mere, mens det for en indvandrer i gennemsnit sænker sandsynligheden med 32%. For efterkommere har knoglesygdommene ikke nogen statistisk signifikant betydning. For danskere har hjerte-kar-syg-

dommen ingen betydning, mens de for indvandrere og efterkommere sænker chancen med gennemsnitligt 25-29%.

Vi vurderer, at disse resultater for, hvordan helbred påvirker lønnen, er ganske store. Imidlertid har tidligere undersøgelser påpeget, at en persons nuværende helbred er korreleret med hans historiske tilknytning til arbejdsmarkedet. Flere undersøgelser finder således, at ledighed har en direkte negativ indvirkning på både det fysiske og det mentale helbred (se fx resultater på danske data i Andersen (2010) samt Browning, Danø & Heinesen (2006)). Lav eller slet ingen beskæftigelse kan påvirke helbredet gennem bl.a. lav fysisk aktivitet, lav følelse af tilhørighed og nytte, men også gennem en mere risikobetonet adfærd, der fx kan komme til udtryk ved et højere forbrug af alkohol (Jahoda 1981). Dette forhold gør, at et individs nuværende helbred kan være en funktion af personens historiske arbejdsmarkedstilknytning. Når vi i tabellen ikke kontrollerer for den historiske arbejdsmarkedstilknytning, så risikerer vi at overvurdere den faktiske betydning af helbredet for den nuværende arbejdsmarkedstilknytning, netop fordi tidligere arbejdsløshed i sig selv kan påvirke helbredet. Sagt på en anden måde, betyder det, at hvis personer med dårligt helbred i højere grad var arbejdsløse tidligere, og den tidligere arbejdsløshed har haft negativ indvirkning på både personernes beskæftigelsespotentiale og deres helbred, så vil resultatet for helbred være overvurderet, når vi ikke kontrollerer for tidligere arbejdsmarkedstilknytning. Helt generelt kan vores resultater være farvet af mange uobserverede variable (fx motivation, evner, misbrug), som potentielt både kan være korreleret med analysens forklarende variable og med udfaldsvariablen. Vi styrker derfor troværdigheden af vores estimater, når vi i de følgende analyser inkluderer proxyer for tidligere arbejdsmarkedstilknytning i analysen.

Vi har nu diskuteret, hvordan de grundlæggende forklarende variable påvirker sandsynligheden for at tjene 125.000 kr. eller derover. Men hvordan påvirker de samme forklarende variable sandsynligheden i den modsatte ende af lønfordelingen, dvs. risikoen for slet ikke at have nogen lønindtægt i første halvår af 2011?

For danskere så vi, at højere alder var både en statistisk og økonomisk signifikant risikofaktor, men når vi i stedet for løn over 125.000 kr. ser på risikoen for ikke at have nogen lønindkomst, så er alder i ringere grad en risikofaktor. For indvandrerne er der forskel på, om personen er en mand eller en kvinde. Yngre kvinder har højere risiko for ikke at have nogen løn, mens det omvendte gælder for unge mænd. Hvis yngre indvandrerkvinder i højere grad er hjemmegående og modtager en offentlig overførselsindkomst, eller de er forsørget af familien, så vil det kunne forklare mønstret. Ældre kvinder og mænd med indvandrerbaggrund har lavere risiko for ikke at have nogen lønindkomst end 35-39-årige. Vi kan ikke tolke på resultaterne for efterkommere, da de er statistisk insignifikante.

Resultaterne viser også, at mænd, uanset om de er indvandrere eller danskere, har højere risiko for ikke at have nogen lønindkomst. Det er overraskende, da tabellens venstre panel viste, at mænd havde højere chance for at tjene 125.000 kr. eller over. Det betyder, at mænd har lavere risiko for at tjene mellem 1 kr. og 124.999 kr. Mandlige efterkommere har lavere risiko for ikke at have nogen lønindkomst. Med hensyn til personer, der ikke er gift, så vi tidligere, at de mandlige indvandrere kun havde en halvt så stor chance for at tjene over 125.000 kr. Resultaterne for en nul-lønindkomst viser modsat, at indvandrermand, der ikke

er gift, har en næsten dobbelt så høj risiko for ikke at tjene noget i første halvår af 2011. Danskere og indvandrerkvinder med fire eller flere hjemmeboende børn har på samme måde kun en halvt så stor chance og en dobbelt så høj risiko for henholdsvis at tjene 125.000 kr. og ikke at tjene noget.

Vi så, at uddannelse var en meget vigtig faktor for, om en person tjente 125.000 kr. eller derover, men også at sammenhængen var mere entydig for danskere end for indvandrere. Resultaterne for danskere viser også klart, at højere uddannede har langt lavere risiko for ikke at have nogen lønindkomst. For indvandrere er resultaterne mere blandede. For indvandrer-mænd er det kun en lang videregående uddannelse, der giver en signifikant lavere risiko for ikke at tjene noget. For kvinder sænker en erhvervsfaglig uddannelse risikoen, mens både en gymnasial uddannelsesbaggrund og en kort videregående uddannelse øger risikoen for ikke at tjene noget. For kvinder bemærker vi også, at den store gruppe med manglende uddannelsesoplysninger har en signifikant større risiko for ikke at have nogen lønindkomst. For efterkommere virker uddannelse på samme måde som for danskerne – personer med en højere uddannelse har lavere risiko for ikke at tjene noget.

Flygtninge har i gennemsnit en cirka dobbelt så høj risiko for ikke at have nogen lønindkomst, men det er værre for kvinderne end for mændene. Flygtede kvinder har en 69% højere risiko for ikke at tjene noget, mens flygtede mænd har en 27% højere risiko. Personer med lav risiko for at opleve sprogproblemer har en 64% lavere risiko for ikke at tjene noget sammenlignet med personer, der har en lav til mellemhøj risiko for sprogproblemer (dvs. personer, hvor risikoen for sprogproblemer er i området 25-50%). Personer med en mellemhøj til høj risiko for at opleve sprogproblemer (dvs. 50-75%) har 82% højere risiko for ikke at tjene noget, mens personer med en estimeret høj sandsynlighed for sprogproblemer (dvs. over 75%), har en risiko for ikke at tjene noget, der er knap en faktor fem højere. Selvvurderede sproglige kompetencer ser altså ud til at være en særdeles vigtig risikofaktor, når vi ser på lønindkomst, men som diskuteret ovenfor, så kan den måde, en person selv vurderer sit sprog på, sagtens være under indflydelse af, hvordan han eller hun klarer sig på arbejdsmarkedet. Blandt andet derfor er sprogvariablen endogen i lønligningerne, hvorfor vi må tolke resultaterne med forsigtighed, da de sandsynligvis er overvurderede.

Ligesom i ligningen for løn på 125.000 kr. finder vi også her en både statistisk og økonomisk meget stærk sammenhæng mellem psykisk sygdom og lav arbejdsmarkedstilknytning. Som minimum viser resultaterne, at psykisk sygdom fordobler risikoen for ikke at have en lønindkomst, men i værste tilfælde øger det risikoen med godt 170% for mandlige indvandre-re. Knoglesygdomme og hjerte-kar-sygdomme er også en risikofaktor, men i langt mindre grad. I værste tilfælde øger en skelet-, muskel- eller knoglesygdom risikoen for ikke at have en løn med 76% for indvandrerkvinder, mens hjerte-kar-sygdomme øger risikoen med 54% for samme gruppe. For danskere har hjerte-kar-sygdomme slet ingen betydning, hvilket bekræfter, hvad vi så ovenfor.

Som sammenfatning på disse indledende resultater baseret på enkle regressionsmodeller med relativt få, men vigtige, forklarende variable, kan vi sige, at de vigtigste faktorer for indvandrerne har været helbred, sprog, og hvorvidt de var flygtet til Danmark. Helbred så kun ud til at være en ubetydelig større risikofaktor for indvandrere, end det var for danskere, men

specielt psykisk sygdom var meget vigtig for begge grupper. Den højeste fuldførte uddannelse havde stor betydning for danskere, men var i langt mindre grad betydningsfuld for indvandrere. For efterkommere var uddannelse vigtigere, end det var for danskerne. Også personer med mange hjemmeboende børn viste sig at være en udsat gruppe, mens danskerne med mange børn i gennemsnit havde lidt højere risiko for lav arbejdsmarkedstilknæytning. Det samme var tilfældet for mænd. Danske mænd havde bedre chance for at tjene 125.000 kr. eller mere end kvinderne, men de havde også større risiko for ikke at have nogen lønindkomst. For indvandrer-mænd var fortegnene på disse risici de samme, men sandsynligheden var tættere på kvinders, end det var tilfældet for de danske mænd. Vi fandt, at højere alder øgede chancen for at tjene 125.000 kr. eller mere, men for danskere var det ikke lige så vigtigt, når vi så på risikoen for ikke at have nogen løn. For indvandrere var højere alder dog også en signifikant og økonomisk betydningsfuld risikofaktor for ikke at have nogen lønindkomst.

Inden vi går videre til analysen af de fuldt specificerede modeller, så skal vi nævne et vigtigt forhold vedrørende frafaldet i datagrundlaget mellem 2009 og 2011. For indvandrere så vi i tabellerne i dataafsnittet ovenfor, at der var knap 173.000 observationer i vores grundlæggende datasæt for 2009. Imidlertid er regressionerne for løn i 2011 udført på baggrund af knap 168.000 indvandrere, hvilket betyder, at der er et frafald på ca. 5.000 personer mellem ultimo 2009 og 2011. I mindre udstrækning gør det samme sig gældende for danskere og efterkommere. Fra-faldet opstår primært på grund af udvandring, men også dødsfald i perioden. Hos danskerne er fra-faldet på 0,4%, mens det hos indvandrere er på 3,0%, mod 1,2% hos efterkommerne. Fra-faldet er naturligt, fordi indvandrere gennemsnitligt er en mere mobil befolkningsgruppe, og problemet i en registerbaseret undersøgelse er langt lavere, end det ville have været i en spørgeskemaundersøgelse. Udfordringen er, at fra-faldet ikke sker tilfældigt. De udvandrede personer er hovedsageligt unge mænd, som har opholdt sig i Danmark under fem år. Vi har undersøgt, om fra-faldet har væsentlig betydning for resultaterne ved at køre modellerne med de udvandrede personer og gøre forskellige antagelser om deres hypotetiske indkomst. Konklusionen er, at resultaternes overordnede tendenser er relativt ufølsomme over for, hvilke antagelser man gør sig om størrelsen af den hypotetiske indkomst for de personer, der er udvandret.

Vi har nu gennemgået resultaterne for to modeller for lønindkomst med et begrænset sæt af forklarende variable. Disse resultater giver os et god udgangspunkt for at undersøge, hvordan de enkelte estimater ændrer sig, når vi inkluderer flere forklaringsfaktorer. Når vi inkluderer flere forklarende variable, bliver modellen bedre til at forudsige, om en person tjener fx 125.000 kr. eller mere. Vores formål er imidlertid ikke at prædikere chancen for at have en vis indkomst, men derimod at undersøge, hvorvidt et karakteristikum sænker eller øger sandsynligheden for at tjene 125.000 kr. Hvis vi inkluderer en ny variabel, som både påvirker lønnen, og med en variabel, som allerede er i vores model, så vil den estimerede koefficient på den oprindelige variabel ændre værdi. I en effektanalyse, hvor man er interesseret i at afdække årsagssammenhænge, vil man være interesseret i at kontrollere for alle forhold, der samtidig både påvirker lønnen og den variabel, som den kausale interesse retter sig imod. Konkret betyder det, at hvis man som i det eksempel, vi brugte ovenfor, var interesseret i den kausale effekt af psykisk sygdom på et arbejdsmarkedsudfald, så ville man kontrollere for alt, hvad

der både påvirker arbejdsmarkedsudfaldet og risikoen for at få en psykisk sygdom. Eksempler kunne inkludere misbrug af alkohol og euforiserende stoffer samt tidligere arbejdsmarkedstilknytning. Vores registerdata kan i nogen udstrækning bruges til at kontrollere for sådanne forhold, men selv når vi har inkluderet alle forklarende variable, der er i vores datasæt, vil der stadig være uobserverede faktorer, der samtidig kan påvirke fx løn og psykisk helbred, herunder misbrug, som personen ikke allerede er i behandling for.¹² Vores undersøgelse skal tages med dette forbehold in mente, men når vi nedenfor inkluderer flere relevante forklarende variable, så er det med til at reducere betydningen af problemet, simpelthen fordi vi kontrollerer for nogle af de faktorer, der før var uobserverede. Estimatet på forhold som fx psykisk sygdom vil derfor komme tættere på det rigtige, men vi risikerer stadigvæk at overestimere i forhold til den sande effekt.

I de følgende afsnit har vi opdelt analyserne sådan, at vi først ser på sandsynligheden for at tjene 125.000 kr. eller mere; bagefter ser vi på risikoen for ikke at have nogen lønindkomst i første halvår af 2011. Efter disse analyser for 2011, analyserer vi samme model, men blot for første halvår 2008, for at undersøge, om tiden før finanskrisen producerer væsentligt anderledes resultater. Afslutningsvis ser vi på overgangen til førtidspension mellem 2007 og 2012.

¹² I forskningssammenhæng kunne man løse dette problem ved hjælp af et randomiseret forsøg, hvor fordelingen af de uobserverede karakteristika var ens i gruppen af personer med og uden mentale helbredsproblemer (selvom det er svært at tænke sig, hvordan sådan et forsøg skulle kunne designes på en etisk forsvarlig måde).

Tabel 4.1 Resultater for model med få forklarende variable for (1) løn på 125.000 kr. eller mere og (2) ingen lønindkomst

	Lønindkomst på 125.000 kr. eller mere					Ingen lønindkomst				
	Danskere	Indvandrere			Efterkommere	Danskere	Indvandrere			Efterkommere
		Alle	Mænd	Kvinder			Alle	Mænd	Kvinder	
25-29 år	0,448**	0,604**	0,725**	0,482**	0,501**	1,030**	1,056**	0,782**	1,362**	0,966
30-34 år	0,808**	0,910**	1,054+	0,800**	0,839**	1,030**	1,015	0,803**	1,197**	1,029
40-44 år	1,157**	1,301**	1,111**	1,488**	1,179	0,984*	0,853**	0,958	0,780**	0,830
45-49 år	1,256**	1,664**	1,379**	1,917**	1,678+	0,963**	0,687**	0,813**	0,618**	0,754
50-54 år	1,327**	1,360**	1,133**	1,515**	1,468	0,925**	0,864**	0,977	0,826**	0,970
55-59 år	1,247**	1,535**	1,359**	1,511**	1,129	0,996	0,767**	0,831**	0,806**	1,060
Mand	1,302**	1,177**			1,522**	1,181**	1,051**			0,893**
Ikke gift	0,695**	0,598**	0,522**	0,667**	0,715**	1,522**	1,839**	1,836**	1,866**	1,326**
Ingen hjemmeboende børn	0,762**	0,903**	0,759**	1,137**	0,908+	1,158**	1,061**	1,281**	0,881**	1,014
1 barn	0,880**	1,020	1,031	1,008	0,921	1,111**	0,939**	0,948+	0,932**	1,095
3 børn	0,804**	0,795**	0,861**	0,742**	0,742**	1,269**	1,261**	1,134**	1,346**	1,296**
4 børn eller mere	0,502**	0,736**	0,876**	0,594**	0,414**	1,961**	1,371**	1,150**	1,591**	1,666**
Gymnasial uddannelse	1,102**	0,787**	0,856**	0,737**	1,303**	0,720**	1,157**	1,030	1,252**	0,574**
Erhvervsfaglig uddannelse	1,928**	1,153**	0,990	1,305**	3,148**	0,534**	0,959*	1,037	0,911**	0,382**
Kort videregående uddannelse	2,334**	0,967	0,999	0,891**	2,423**	0,478**	1,142**	1,029	1,308**	0,410**
Mellemlang videregående uddannelse	4,908**	1,318**	1,222**	1,411**	6,864**	0,240**	0,983	0,986	1,002	0,181**
Lang videregående uddannelse	2,847**	1,416**	1,412**	1,415**	4,107**	0,411**	0,944*	0,845**	1,074+	0,263**
Manglende uddannelsesoplysninger	0,502**	1,001	1,069*	0,888**	0,675**	1,815**	1,119**	0,938*	1,331**	1,290*
Flygtning eller vurderet flygtning		0,755**	0,837**	0,707**			1,494**	1,266**	1,688**	
Lav risiko for sprogproblemer		2,882**	3,073**	2,661**			0,356**	0,359**	0,349**	
Mellemhøj risiko for sprogproblemer		0,515**	0,507**	0,497**			1,818**	1,781**	2,004**	
Høj risiko for sprogproblemer		0,185**	0,232**	0,172**			4,812**	3,257**	5,732**	
Psykisk sygdom	0,456**	0,412**	0,360**	0,442**	0,505**	2,490**	2,608**	2,720**	2,643**	2,037**
Skelet-, knogle- eller muskelsygdomme	0,768**	0,680**	0,672**	0,657**	0,896	1,385**	1,621**	1,552**	1,755**	1,256**
Hjerte-kar-sygdom	0,996	0,808**	0,862**	0,743**	0,747*	0,993	1,347**	1,204**	1,535**	1,290*
Antal personer	2.128.836	167.801	78.766	89.035	13.651	2.128.836	167.801	78.766	89.035	13.651

Referencepersonen er en gift, 35-39-årig kvinde med to børn og en grundskoleuddannelse. Hvis hun er indvander, så er hun ikke flygtning, hun har boet i Danmark i 5-9 år og har lav til mellemhøj risiko for sprogproblemer. I tillæg til de viste variable, så kontrolleres der også for, om personen bor i en af landets fire største byer og for opholdstid i Danmark. Modellernes evne til at forklare variationen i udfaldsvariablene, målt ved de såkaldte pseudo R², er (0,11; 0,17; 0,14; 0,20; 0,12) og (0,07; 0,14; 0,11; 0,18; 0,07) henholdsvis for modellerne for løn over 125.000 kr. og ingen lønindkomst. '+' indikerer, at odds-ratioen er signifikant forskellig fra 1 på et 10%-s niveau, mens '**' og '**' angiver signifikans på hhv. 5% og 1%-s niveau.

4.2 Udvidet model for lønindkomst over 125.000 kr.

I afsnittet ovenfor tolkede vi på resultaterne fra to regressioner for lønindkomst, men vi inkluderede kun nogle få grundlæggende demografiske, socioøkonomiske og helbredsmæssigt forklarende variable i analyserne. I dette afsnit gentager vi analysen af, hvilke faktorer der er bestemmende for, om en person har en lønindkomst på 125.000 kr. eller mere i 2011, men vi inkluderer langt flere kontrolvariable. De ekstra kontrolvariable omfatter et langt mere detaljeret sæt af helbredsvariable, historisk og nuværende arbejdsmarkedstilknytning samt oprindelsesland for indvandrere.

Fordi regressionerne nu inkluderer flere variable, bliver listen af resultater også meget længere. Vi har derfor valgt at opdele resultaterne i separate tabeller efter variabelkategori. I Tabel 4.2 viser vi odds-ratio-estimerne for demografiske og socioøkonomiske variable, mens Tabel 4.3 viser resultaterne for branchetilhørsforhold og typen af beskæftigelse. I Tabel 4.4 viser vi resultaterne for de indvandrespecifikke variable – dvs. opholdstid, oprindelsesland, opholdsgrundlag og sprog – mens Tabel 4.5 viser, hvordan vores samlede sæt helbredsmæssige indikatorer påvirker evnen til at opretholde det givne lønniveau. Som et alternativ til den meget detaljerede helbredsinddeling viser Tabel 4.6 den enklere opdeling, som vi brugte ovenfor, med kun tre typer af helbredsproblemer. I det følgende diskuterer vi resultaterne og de forbehold, der måtte være og sammenligner med det, vi tidligere har set.

Resultaterne i Tabel 4.2 viser, ikke overraskende, at unge under 30 år har lavere sandsynlighed for at have en høj indkomst end de 35-39-årige. Eksempelvis viser tabellens første række, at odds-ratioen er 0,671 gange lavere for danskere i alderen 25-29 år end danskere i 35-39-årsalderen. Det svarer til knap 33% (dvs. 1-0,671). Personer i alderen 30-34 og 40-44 år har en sandsynlighed tæt på referencegruppen af 35-39-årige, idet odds-ratioen ikke er signifikant forskellig fra en. I modsætning til, hvad vi så ovenfor, så falder sandsynligheden for at tjene over 125.000 kr., når vi ser på gradvist ældre befolkningsgrupper. Dette gælder både for indvandrere og danskere. Denne ændring skyldes, at vi nu kontrollerer for langt flere variable – herunder erhvervserfaring – end vi gjorde ovenfor. Resultaterne for efterkommere viser samme mønster, men er mere statistisk usikre. Vi konkluderer derfor, at når der er redegjort for erhvervserfaring, så er højere alder en risikofaktor for arbejdsmarkedstilknytningen for både danskere og indvandrere. Imidlertid betyder det ikke nødvendigvis, at ældre har det svært på arbejdsmarkedet. Det kan blot være udtryk for, at ældre har højere præferencer for fritid, og at de derfor eksempelvis kun arbejder på halv tid.

Tabellen viser, at ugifte indvandermænd har en 20% lavere sandsynlighed end gifte for at tjene 125.000 kr. eller mere, mens sandsynligheden for ugifte kvinder ikke adskiller sig fra gifte kvinder. Efter vi har inkluderet flere kontrolvariable, viser civilstand sig altså at være en langt mindre risikofaktor, end vi så ovenfor. Forskellen mellem mænd og kvinder kan skyldes, at indvandermændene påtager sig en forsørgerrolle, når de gifter sig, mens kvindernes beskæftigelse ikke afhænger af, om de har indgået ægteskab eller ej. Resultaterne, vi præsenterer i næste afsnit, viser, at kvinder, der ikke er gift, har en højere risiko for slet ikke at have nogen lønindkomst.

Personer uden hjemmeboende børn og personer med mange hjemmeboende børn har ringere chance for at tjene over 125.000 kr., og i modsætning til tidligere øges chancen ikke længere for kvinder uden hjemmeboende børn.

Tabel 4.2 Logistisk regression for lønindkomst i 2011; demografiske og socioøkonomiske variable

	Danskere	Indvandrere			Efter-kommere
		Alle	Mænd	Kvinder	
25-29 år	0,671**	0,643**	0,671**	0,612**	0,660**
30-34 år	1,079**	0,951*	1,003	0,898**	1,056
40-44 år	0,851**	1,021	0,881**	1,158**	0,854
45-49 år	0,678**	0,893**	0,763**	1,028	0,964
50-54 år	0,604**	0,807**	0,681**	0,940	1,107
55-59 år	0,532**	0,659**	0,560**	0,745**	0,665
Mand	1,664**	1,679**			1,697**
Ikke gift	0,886**	0,922**	0,801**	1,043	0,774**
Ingen hjemmeboende børn	0,708**	0,770**	0,740**	0,828**	0,834**
1 barn	0,900**	0,920**	0,960	0,872**	0,936
3 børn	0,955**	0,924**	0,977	0,907**	0,837+
4 børn eller flere	0,776**	0,790**	0,837**	0,770**	0,582**
Årsledighedsgrad 1-19%	0,539**	0,545**	0,563**	0,526**	0,464**
Årsledighedsgrad 20% og derover	0,384**	0,381**	0,433**	0,323**	0,393**
Erhvervsgrad mellem 33 og 65%	2,288**	2,623**	2,499**	2,882**	2,942**
Erhvervsgrad 66% og derover	5,671**	4,965**	4,885**	5,599**	6,227**
Har modtaget sygedagpenge	0,753**	0,900**	0,933*	0,847**	0,959
København	0,924**	1,065*	1,037	1,098*	0,961
Vestegnskommuner	1,138**	1,116**	1,125**	1,105**	1,121
Aarhus	0,822**	0,886**	0,925+	0,828**	1,128
Aalborg	0,987	1,000	0,906	1,094	0,950
Odense	0,852**	0,859**	0,824**	0,907	0,745*
*Indv.andel i kom. (kvintil)	1,029**	1,050**	1,022	1,076**	0,991
*Lokal ledighed	0,972**	0,966**	0,978*	0,956**	1,008
Gymnasial uddannelse	0,806**	0,946*	0,995	0,927*	0,863+
Erhvervsfaglig uddannelse	1,464**	1,452**	1,227**	1,679**	2,061**
Kort videregående uddannelse	1,648**	1,332**	1,356**	1,282**	1,423**
Mellemlang videregående uddannelse	2,222**	1,579**	1,502**	1,700**	2,955**
Lang videregående uddannelse	1,751**	1,445**	1,403**	1,521**	1,919**
Manglende uddannelsesoplysninger	0,773**	0,997	0,966	1,035	0,855
Lønmodtagere på højeste niveau	1,241**	1,272**	1,261**	1,290**	1,469**
Øvrige lønmodtagere	0,755**	0,744**	0,617**	0,932	0,905
Selvstændig	0,040**	0,048**	0,059**	0,042**	0,016**
Ledige (i uge 48)	0,559**	0,587**	0,627+	0,504*	0,098**
Syge- og barselsdagpenge	0,354**	0,271**	0,165**	0,304**	0,058**
Aktivering og rehabilitering	0,186**	0,234**	0,292**	0,158**	0,043**
Kontanthjælp og integrationsydelse	0,036**	0,054**	0,071**	0,036**	0,009**
Øvrige uden for arbejdsstyrken	0,192**	0,141**	0,214**	0,080**	0,035**
Under uddannelse	0,208**	0,188**	0,179**	0,163**	0,035**
Antal personer	2.128.836	167.801	78.766	89.035	13.651

Referencepersonen er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i årets løb. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet, har boet i Danmark i 5-9 år og har 25-50%’s risiko for at opleve sproglige problemer. ‘+’ indikerer, at odds-ratioen er signifikant forskellig fra 1 på et 10%’s niveau, mens ‘**’ og ‘***’ angiver signifikans på hhv. 5% og 1%’s niveau.

De nye kontrolvariable, vi har inkluderet, omfatter mål for erhvervserfaring, arbejdsløshed og sygdom. Referencepersonen har en erhvervsgrad under 33%, hvilket betyder, at hun ikke har nogen eller kun lidt erhvervserfaring siden 1980. Samtidig har hun ikke været ledig eller på sygedagpenge i 2009. Resultaterne for de nye variable viser, at sammenlignet med referencepersonen har en person med mellemhøj erhvervsgrad over dobbelt så stor sandsynlighed for at have en løn på 125.000 kr. eller mere, mens en person med en høj erhvervsgrad, og dermed lang tids erhvervserfaring har en fem til seks gange så stor sandsynlighed. Erhvervserfaring er altså – ikke overraskende – en særdels stærk prædikator for lønnen for både danskere, indvandrere og efterkommere. Personer med en høj ledighed i 2009 eller personer, der har modtaget sygedagpenge, har også lavere chance for at tjene 125.000 kr. eller mere i første halvår af 2011.

Vi kan konstatere, at indvandrere, som bor i kommuner med en højere indvandrerandel, har en lidt højere sandsynlighed for at tjene 125.000 kr. eller mere, men estimatet er ikke væsentligt anderledes for indvandrere end for danskere. Resultatet understøtter derfor kun i begrænset omfang hypotesen om, at et godt socialt netværk har betydning. Indvandrere, der bor i hovedstadsområdet, har også bedre chancer for at have en rimelig løn, mens det trækker ned at bo i de største provinskommuner. Hvis en person derimod bor i en kommune, hvor ledigheden er højere, reduceres chancen for at tjene 125.000 kr. Når vi kontrollerer for lokal arbejdsløshed, så indikerer det positive resultat for hovedstadskommunerne for indvandrere, at der er forhold forbundet med at bo i hovedstadsområdet, som hjælper indvandrerne til en bedre arbejdsmarkedstilknytning, men vi kan ikke afvise, at resultaterne stadig er påvirket af uobserverede underliggende sammenhænge.

Når vi ser på estimaterne for de forskellige uddannelseskategorier, får vi overordnet stadig indtryk af, at bedre uddannede klarer sig bedre, selvom estimaterne er lavere, end det vi så ovenfor. Selvom disse resultater ikke nødvendigvis er årsagssammenhænge, så bekræfter de alligevel den eksisterende viden: En erhvervsfaglig eller videregående uddannelse øger chancen for at have en god arbejdsmarkedstilknytning ganske betragteligt i forhold til kun at have en grundskoleuddannelse. Estimaterne er lavere for indvandrergruppen end for danskerne, men for efterkommerne øger uddannelse sandsynligheden for at tjene 125.000 kr. eller mere end for de øvrige grupper.

Eksempelvis mere end fordobler en faglært uddannelse chancen for at have denne indkomst for efterkommere, mens det for mandlige indvandrere øger chancen med 23%. For indvandrerkvinderne indikerer tallene, at gevinsten ved at tage en lang eller mellemlang uddannelse er marginalt højere end for mændene. Vi kan ikke sige, om det i realiteten blot er udtryk for selvselektion, sådan at de kvinder, der færdiggør en uddannelse, er mere motive-rede med bedre evner, hvor mændene måske er en bredere sammensat gruppe. Resultaterne viser også, at en gymnasial uddannelse ikke er tilstrækkelig til at differentiere sig fra personer med blot en grundskoleuddannelse; faktisk viser resultaterne, at personer, der højst har en gymnasial uddannelse, har mindre chance for at tjene 125.000 kr. eller mere. Endelig kan vi bemærke, at den femtedel af indvandrerne, som havde manglende uddannelsesoplysninger, ikke adskiller sig signifikant fra personer med blot en grundskoleuddannelse.

Vi så ovenfor, at erhvervserfaring var meget vigtig, så det er ikke overraskende, at lønmodtagere på højeste niveau – der også må formodes at have en historisk god tilknytning til arbejdsmarkedet – har en større sandsynlighed for at have en løn på 125.000 kr. eller mere sammenlignet med lønmodtagere på mellem- og grundniveau. Lønmodtagere på laveste niveau har en ca. 25% lavere sandsynlighed. De selvstændige har typisk ikke en lønmodtagerindkomst, hvorfor resultatet for denne gruppe ikke bør fortolkes, men blot opfattes som en kontrolvariabel; vi undersøger dette forhold nærmere i Bilag 2: Separat analyse for selvstændiges arbejdsmarkedstilknytning og konkluderer, at de selvstændige ikke umiddelbart adskiller sig, når man ser på nettooverskuddet fra driften af deres virksomheder. Det er heller ikke overraskende, at ledige, syge, personer under uddannelse og de mere marginaliserede i det hele taget har en langt lavere sandsynlighed for at opretholde en løn på 125.000 kr. eller mere. Overordnet kan man sige, at nuværende arbejdsmarkedstilknytning er en særdeles stærk prædiktør for fremtidig arbejdsmarkedstilknytning.

Tabel 4.3 viser, hvordan beskæftigelse i forskellige brancher og forskellige typer hovedbeskæftigelse påvirker sandsynligheden for at have en løn over det definerede niveau. Alle sandsynlighederne måles op imod det mest almindelige, som er 'uklassificeret', hvilket i praksis er ensbetydende med, at personen har en lavere arbejdsmarkedstilknytning. Derfor er det ikke underligt, at stort set alle odds-ratioer er større end en.¹³ De relative størrelser af estimerne er imidlertid mere informative end deres niveau, men det er vanskeligt at fortolke betydningen i og med, at referencegruppen kan være meget forskelligt sammensat, men tendensen er, at brancher, der i højere grad benytter specialiseret og højtuddannet arbejdskraft, også er de brancher, hvor de ansatte har højere sandsynlighed for at tjene 125.000 kr. eller derover i første halvår af 2011.

¹³ Brancheoplysninger for efterkommer er udeladt i den endelige model, da de gjorde modellen ustabil på grund af den lille population.

Tabel 4.3 Logistisk regression for lønindkomst i første halvår 2011; branche og typen af beskæftigelse

	Danskere	Indvandrere			Efterkom- mere
		Alle	Mænd	Kvinder	
Landbrug	1,530**	3,289**	4,131**	1,838+	-
Industri	2,109**	2,490**	3,166**	1,714+	-
Byggeri og anlæg	1,117+	1,365	1,706*	0,785	-
Handel og transport	1,425**	1,586*	2,243**	0,895	-
Kommunikation og it	1,787**	3,020**	4,089**	1,868+	-
Finans og forsikring	3,424**	3,762**	4,292**	2,656**	-
Ejendomshandel og udlejning	2,202**	2,891**	5,188**	1,378	-
Erhvervsservice (fx rengøring)	1,463**	1,525*	1,971**	0,996	-
Offentlig sektor	1,680**	2,319**	3,242**	1,450	-
Kultur og fritidsaktiviteter	1,207**	1,813**	1,902*	1,345	-
Militært arbejde	3,787**	3,415**	3,013**	2,299	5,599**
Ledelse på øverste niveau	2,857**	3,882**	3,191**	5,266**	3,203**
Højt specialiseret arbejde	3,075**	4,664**	3,891**	5,637**	3,683**
Arbejde, der forudsætter viden på mellemniveau	2,453**	3,684**	2,923**	4,504**	3,217**
Kontorarbejde	1,614**	2,479**	2,007**	3,022**	2,516**
Detailsalg, service- og omsorgsarbejde	1,089**	1,469**	1,212**	1,771**	1,543**
Arbejde i de primære erhverv	0,804**	1,298	1,243	1,580	0,977
Håndværkspræget arbejde	1,402**	1,861**	1,745**	1,713**	1,853**
Maskin-, transport- og anlægsarbejde	1,532**	2,855**	2,547**	3,138**	2,074**
Andet arbejde	1,211**	2,029**	2,077**	1,904**	1,777**
Antal personer	2.128.836	167.801	78.766	89.035	13.651

Baseline er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i 2009. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet, har boet i Danmark i 5-9 år, og har 25-50%’s risiko for at opleve sproglige problemer. ‘+’ indikerer, at odds-ratioen er signifikant forskellig fra en på et 10%’s niveau, mens ‘*’ og ‘**’ angiver signifikans på hhv. 5% og 1%’s niveau. ‘-’ indikerer, at variablen ikke er medtaget i analysen.

Tabel 4.4 viser resultaterne for de indvandrerspecifikke variable. Disse variable omfatter opholdstiden, oprindelseslandet, opholdsgrundlag samt de estimerede sprogkompetencer. For efterkommerne eksisterer der kun oplysninger om oprindelsesland – de har boet i Danmark hele livet og antages ikke at have sprogproblemer. Resultaterne viser, at de ikke-vestlige indvandrere har større chance for at tjene over 125.000 kr. des længere tid de har opholdt sig i landet, og det gælder i lidt højere grad for kvinder. Flygtningestatus og indikatoren for sprog-vanskeligheder har ikke nogen betydning, når vi samtidig kontrollerer grundigt for oprindelsesland såvel som historisk og nuværende arbejdsmarkedsstatus.

I den indledende analyse og i Bilag 3: Alternative modelspecifikationer så vi det stik modsatte. Her var sprogkompetencer en meget vigtig faktor for arbejdsmarkedstilknytningen, men det gælder altså kun, indtil vi kontrollerer grundigt for blandt andet den historiske arbejdsmarkedstilknytning. I afsnittet ovenfor, hvor vi ikke kontrollerer for så mange faktorer, så vi, at gode sprogkompetencer øgede sandsynligheden for at tjene 125.000 kr. eller mere 188% i forhold til en person med mellemgode sprogkompetencer, mens personer med de

dårligste selv vurderede sprogkompetencer havde en sandsynlighed, der var 81% lavere. Sprog er altså kun vigtigt for sandsynligheden for lønnen, hvis man ikke tager højde for personens nuværende og tidligere tilknytning til arbejdsmarkedet samt hans oprindelsesland. Gør man derimod det, så har sprog ikke nogen signifikant betydning. Lidt senere ser vi igen på, hvad der forklarer risikoen for at have en lønindkomst på nul kroner. Her vil vi se, at dårlige sprogkompetencer øger risikoen samtidig med, at gode sprogkompetencer sænker risikoen for ikke at have nogen lønindkomst. Denne sammenhæng og forskellen mellem resultaterne kan både være udtryk for, at det er de *selv vurderede* sprogkompetencer, vi måler på, hvorfor nuværende beskæftigelse kan være afgørende for, hvorvidt en indvandrer vurderer, om hans sprog er godt eller dårligt. Forskellen i betydningen af sprog, alt efter om vi kontrollerer for tidligere arbejdsmarkedstilknytning eller ej, kan også skyldes, at personer med dårlige sprogkompetencer kun sjældent tjener 125.000 kr. eller mere (se resultaterne i Bilag 4), hvorfor sprog helt mister sin forklaringskraft, når vi inkluderer en kontrol for arbejdsmarkedstilknytningen. Resultaterne kan altså være et udtryk for, at sprog er en vigtig risikofaktor for personer, der ikke har en tilknytning til arbejdsmarkedet, mens sprog ikke er en risikofaktor for personer, der har en god tilknytning til arbejdsmarkedet.

Kvinder, der er indvandret fra bl.a. Somalia, Libanon, Irak, Pakistan, Syrien og Afghanistan, adskiller sig fra tyrkiske kvinder, som er referencegruppen, ved at have en signifikant lavere sandsynlighed for at have en indkomst over 125.000 kr. Mændene fra disse lande har en sandsynlighed, der enten er højere eller ikke adskiller sig signifikant fra tyrkernes. Mænd og kvinder fra bl.a. Iran, lande i det tidligere Jugoslavien, Rusland, Rumænien samt visse asiatiske lande adskiller sig ved at have signifikant større chance for at tjene over 125.000 kr. Disse mønstre kan skyldes, at indvandrere fra nogle lande medbringer erhvervskompetencer og en arbejdskultur, der foranlediger en let integrering på det danske arbejdsmarked.

Ovenfor så vi, at det var en signifikant risikofaktor, hvis en indvandrer var flygtet til landet. Som vi har set med sprogvariablen, så gør det ikke nogen signifikant forskel, om en indvandrer er flygtet eller ej, under forudsætning af, at vi kontrollerer for arbejdsmarkedsstatus og oprindelsesland. I og med, at flygtninge ofte kommer fra nogle specifikke lande, er det måske ikke overraskende, at landevariablene trækker forklaringskraft fra flygtningeindikatoren.

Tabel 4.4 Logistisk regression for lønindkomst i 2011; indvandrerspecifikke variable

	Danskere	Indvandrere			Efter- kommere
		Alle	Mænd	Kvinder	
Højst 4 år siden indrejse		0,863**	0,922+	0,793**	
10-14 år siden indrejse		1,098**	1,085*	1,101**	
15-19 år siden indrejse		1,113**	1,063+	1,142**	
Minimum 20 år siden indrejse		1,080**	1,019	1,136**	
Iran		1,188**	1,310**	1,130+	0,782
Irak		0,871**	0,969	0,790**	1,904
Vietnam		1,245**	1,218**	1,312**	0,794+
Libanon		0,941	1,138*	0,748**	0,559*
Pakistan		0,974	1,127*	0,779**	0,913
Bosnien		1,580**	1,624**	1,569**	4,219*
Jugoslavien		1,208**	1,230**	1,225**	1,249**
Makedonien		1,561**	1,713**	1,447**	1,418+
Kroatien		1,465**	1,383	1,631*	1,669
Serbien-Montenegro		0,982	1,150	0,844	1,129
Thailand		1,300**	1,424**	1,243**	1,044
Somalia		0,900*	1,077	0,688**	0,773
Sri Lanka		1,285**	1,385**	1,249**	0,476*
Filippinerne		1,256**	1,317**	1,228**	1,496*
Marokko		1,268**	1,498**	1,063	1,202+
Afghanistan		0,991	1,090	0,850+	1,319
Kina		0,937	0,813**	1,031	1,378
Ukraine		1,690**	2,016**	1,400**	0,595
Rumænien		1,492**	1,474**	1,507**	2,270
Indien		1,674**	1,648**	1,623**	0,981
Rusland		1,322**	1,433**	1,253**	2,734
Brasilien		1,155*	1,093	1,154	1,214
Bulgarien		1,199*	1,078	1,292*	1,243
Syrien		0,808*	0,995	0,643**	0,675
Ghana		1,241**	1,300*	1,202	3,976+
Nepal		0,498**	0,415**	0,805	0,140*
Øvrige ikke-vestlige lande		1,173**	1,230**	1,120**	0,864
Flygtning eller vurderet flygtning		1,002	1,000	1,001	
Lav risiko for sprogproblemer		1,014	1,026	1,018	
Mellemhøj risiko for sprogproblemer		1,013	0,994	1,030	
Høj risiko for sprogproblemer		0,988	1,041	1,040	
Antal personer	2.128.836	167.801	78.766	89.035	13.651

Baseline er en gift 35-39-årig kvinde med 2 børn. Hun arbejder som ufraglært lønmodtager på mellem eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i 2009. Hvis hun er indvandrer, så er hun indvandret hun fra Tyrkiet, hun har boet i Danmark i 5-9 år, og hun har 25-50% risiko for at opleve sproglige problemer. " + " indikerer at odds-ratioen er signifikant forskellig fra 1 på et 10%-niveau, mens " * " og " ** " angiver signifikans på et hhv. 5%- og 1%-niveau.

Tabel 4.5 viser regressionsresultaterne for samtlige helbredsindikatorer: Køb af visse receptpligtige lægemidler, diagnoser givet i forbindelse med ambulante besøg eller indlæggelse på et sygehus samt lægekontakter under den offentlige sygesikring. De medicintyper, som vi har valgt at inkludere i undersøgelsen, kan alle være udtryk for eksistensen af en sygdom eller et symptom, der i teorien kan have betydning for arbejdsevnen. Disse indikatorer er altså langt mere detaljerede end de tre, som vi omtalte i afsnittet ovenfor. Vi har kørt regressionen med

både det fulde sæt af helbredsindikatorer og det reducerede sæt bestående af tre indikatorer. De sidstnævnte resultater vises i Tabel 4.6.

Mange af helbredsestimaterne er statistisk signifikante for danskerne, mens de i lidt ringere grad er det for indvandrerne. Selvom en del af estimerne er statistisk signifikante, ændrer de kun marginalt på sandsynligheden relativt til referencegruppen. Kun nogle få estimer er signifikante for populationen af efterkommere, hvilket kan skyldes, at observationsantallet er mindre, og at de yngre efterkommere ikke køber medicin så ofte. Overordnet viser resultaterne den forventede negative sammenhæng mellem sygdom og evnen til at opretholde en rimelig lønindtægt.

For lægemidlerne ser vi, at antacid, der bruges mod mavesyrerelaterede forstyrrelser, sænker indvandrernes chance marginalt for at have en indkomst over 125.000 kr. med knap 4%. Insulin, som bruges ved diabetes, sænker også sandsynligheden for at have en rimelig indkomst, men kun for mænd. For danskere er brug af antithrombosemidler, der forhindrer blodpropper, også forbundet med en lavere chance for indkomst over 125.000 kr. Det samme gælder betablokkere, der kan bruges ved en række sygdomme relateret til hjerte-kar-systemet (bl.a. forstyrrelser i hjerterytmen, hjerteanfald og højt blodtryk).

Tabel 4.5 Logistisk regression for lønindkomst i 2011; lægemidler, diagnoser og sygesikringssydelser

	Danskere	Indvandrere			Efterkommere
		Alle	Mænd	Kvinder	
Antacid (A02)	0,944**	0,963+	0,954	0,969	1,000
Insulin (A10)	0,827**	0,811**	0,718**	0,940	1,151
Antithrombose (B01)	0,927**	0,981	0,944	1,052	0,785
Hjerteterapi (C01)	0,942*	0,827+	0,841	0,862	1,314
Diuretica (C03)	1,007	1,008	0,839+	1,038	0,689
Betablokkere (C07)	0,931**	0,975	1,006	0,971	0,871
Calciumantagonister (C08)	0,966**	0,986	0,989	1,008	0,448+
Renin-angiotensin syst. (C09)	1,002	0,978	1,029	0,924	0,743
Lipid-modif. præp. (C10)	1,024*	0,964	1,057	0,872+	0,683
Antireumatica og antiinflam. (M01)	0,979**	0,947**	0,949*	0,945*	1,061
Analgetica (N02)	0,899**	0,855**	0,828**	0,885**	0,874
Pyskoleptica og psykoanaleptica (N05-6)	0,704**	0,720**	0,681**	0,750**	0,749**
Anti-afhængighed (N07B)	0,600**	0,668*	0,582**	1,282	0,069**
KOL-præparater (R03)	1,006	1,070*	1,053	1,074+	0,945
Antihistaminer (R06)	1,021**	1,018	1,020	1,007	0,896
Infektioner (A, B)	0,911**	1,025	1,058	1,017	1,265
Svulster (C00-D48)	0,991	1,022	0,867	1,032	1,524*
Blodsygdomme mv. (D50-D89)	0,945+	0,884	0,915	0,860	1,273
Ernæringssygdomme mv. (E)	0,991	1,042	1,181*	0,989	0,966
Psykiske lidelser (F)	0,603**	0,705**	0,597**	0,835	0,812
Nervesygdomme (G)	0,811**	0,950	1,091	0,873+	0,999
Øre- og øjensygdomme (H)	0,942**	1,001	0,938	1,041	1,094
Kredsløbssygdomme (I)	0,949**	1,016	0,963	1,072	0,878
Åndedrætssygdomme (J)	0,945**	0,975	0,937	1,022	0,917
Fordøjelsessygdomme (K)	0,974**	0,974	0,990	0,961	1,066
Hudsygdomme (L)	0,960**	1,030	1,105	0,977	0,958
Knoglesygdomme mv. (M)	0,901**	0,847**	0,846**	0,846**	1,072
Urinvejssygdomme mv. (N)	1,009	1,057+	1,137*	1,023	0,840+
Svangerskabs., misdannelser mv. (O, P, Q)	1,040**	1,064+	0,788	1,162**	1,346**
Div. sympt. (mavesmerte, besvimelse mv.) (R)	0,936**	0,922**	0,891*	0,939+	0,818*
Ydre påvirkning (ulykke mv.) (S, T, X, Y)	0,954**	0,946+	0,934	0,946	1,085
Undersøgelse (Z00-Z13)	0,960**	0,930**	0,935+	0,933*	0,910
Kontakt ved reproduktion (Z30-Z39)	0,772**	0,726**	1,117	0,749**	0,566**
Øvrige forhold (Z20-Z29, Z40-Z99)	0,979*	0,955	0,979	0,937	1,106
*Alm. læge (antal kontakter)	0,993**	0,996**	0,994*	0,997*	1,003
Tandlæge	1,167**	1,110**	1,133**	1,085**	1,074
Fysioterapi	0,971**	1,005	1,030	0,975	0,904
Kiropraktik	1,064**	1,238**	1,364**	1,125+	1,189
Psykologhjælp	0,839**	0,898**	0,912	0,877**	0,740**
Psykiatri	0,711**	0,673**	0,630**	0,706**	0,722
Hud- og kønssygdomme	1,022**	1,006	1,003	1,013	0,945
Øjenlæge	1,026**	1,008	1,052	0,973	0,928
Ørelæge	1,002	0,970	1,027	0,930+	0,978
Gynækologi	1,123**	1,003	1,365*	0,946	1,126
Reumatologi	0,926**	0,916+	0,990	0,858*	0,777
Intern medicin	1,044+	1,089	1,020	1,140	1,415
Kirurgi (plastik, ortopædisk mv.)	1,040**	1,035	1,043	1,017	1,026
Antal personer	2.128.836	167.801	78.766	89.035	13.651

Baseline er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufraglært lønmodtager på mellem- eller grundniveau, har høj erhvervsgrad og hverken modtaget sygedagpenge eller været ledig. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. '+' indikerer, at odds-ratioen er signifikant forskellig fra 1 på et 10%'s niveau, mens '**' og '**' angiver signifikans på hhv. 5% og 1%'s niveau. Parentesen angiver, hvilken ATC- eller ICD-10 gruppe, der er tale om.

I gruppen af lægemidler, der har indvirkning på nervesystemet, ser vi klart den stærkeste sammenhæng mellem helbred og løn. Analgetica, som virker smertestillende, sænker sandsynligheden for en indkomst over 125.000 kr. med 10-15% i runde tal, mens psykofarmaka såsom antidepressiva og angstdæmpende midler reducerer sandsynligheden med 25-30%. For både danskere, mandlige indvandrere og efterkommere sænker køb af midler mod afhængighed, fx antabus og metadon, sandsynligheden for at have en rimelig indkomst ganske betragteligt.

Lægemidler med virkning på åndedrætssystemet påvirker derimod sandsynligheden en smule i positiv retning.

Diagnoserne fra hospitalskontakter påvirker også overvejende lønnen negativt som forventet. Den stærkeste sammenhæng optræder, som for lægemidlerne, for gruppen af psykiske lidelser. For både indvandrere og danskere er lønnen også signifikant negativt korreleret med diagnoser for knoglesygdomme. Kontakter i forbindelse med graviditet og reproduktion sænker sandsynligheden for at tjene over 125.000 kr., ca. 25% for kvindelige indvandrere og danskere, men endnu mere for efterkommere.

Tabellens sidste gruppe af variable beskriver kontakter til ydere under den offentlige sygesikring. Alle variable er kodet som binære, undtagen kontakter til alment praktiserende læger, der måles som antallet af kontakter. Sidstnævnte variabel må anses som relativt støjfyldt, fordi patienter bruger lægen forskelligt, men fortegnet er negativt som forventet. Besøg hos tandlæge og kirurger påvirker derimod sandsynligheden positivt, men det skyldes formentlig selektion, da der er en høj grad af egenbetaling på disse ydelser. Besøg hos psykolog og psykiater varierer derimod både økonomisk og statistisk signifikant negativt med evnen til at opretholde en rimelig løn. For indvandrere sænker et eller flere besøg hos en psykiater sandsynligheden med ca. 30%. For danskere er det tilsvarende. Psykologbesøg har samme fortegn, men korrelationen er knap så stærk.

Besøg hos øre- og øjenlæger samt specialister i hud- og kønssygdomme er tilsyneladende ikke statistisk forbundet med indkomsten for indvandrerpopulationen. For gynækologbesøg kan læseren se bort fra den positive korrelation mellem mandlige indvandreres indkomst og besøg hos en sådan, da sammenhængen udelukkende er drevet af nogle få såkaldte statistiske outliers. Det sidste forhold, vi vil nævne, er en negativ sammenhæng mellem reumatologiske kontakter og indkomst. For indvandrerkvinder sænker disse kontakter, der kan være relateret til skeletsygdomme såsom gigt, sandsynligheden for at have en lønindkomst på over 125.000 kr. med ca. 15%.

Som nævnt har vi forsøgt at samle nogle bredt definerede store sygdomsgrupper under et. Årsagen er, at de mange helbredsindikatorer kan være korreleret og derfor i nogen grad vil udtrykke det samme. Som beskrevet ovenfor er de tre brede grupper, som vi har defineret, psykisk sygdom, knogle-, muskel- og skeletsygdomme samt hjerte-kar-sygdomme.

Tabel 4.6 Alternativ specifikation af helbredsindikatorer

	Danskere	Indvandrere		
		Alle	Mænd	Kvinder
Psykisk sygdom	0,593**	0,609**	0,551**	0,643**
Knogle-, muskel- og skeletsygdom	0,836**	0,798**	0,793**	0,786**
Hjerte-kar-sygdom	0,909**	0,920**	0,920*	0,933+
Antal personer	2.128.836	167.801	78.766	89.035

Baseline er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i 2009. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. Hun har 25-50%'s risiko for at opleve sproglige problemer. '+' indikerer, at odds-ratioen er signifikant forskellig fra en på et 10%'s niveau, mens '*' og '**' angiver signifikans på hhv. 5% og 1%'s niveau.

For danskerne viser resultaterne i Tabel 4.6, at psykisk sygdom reducerer sandsynligheden for at have en lønindkomst på 125.000 kr. med knap 40%. For indvandrere er det i gennemsnit det samme, men her ser vi også, at mænds indtjeningssevne påvirkes mere end kvinders. Muskel-, skelet- og knoglesygdomme, som fx omfatter ryglidelser, sænker sandsynligheden med 16% for danskere og 20% for indvandrere, mens hjerte-kar-sygdommene sænker risikoen med 8-10%.

Vi har nu diskuteret de relevante resultater fra modellen for lønindkomst over 125.000 kr. i første halvår af 2011. Vi fandt bl.a., at helt unge og ældre indvandrere har det relativt sværere, mens personer med en erhvervs- eller videregående uddannelse klarede sig relativt bedre. Vi fandt også, at nuværende og historisk arbejdsmarkedstilknnytning er særdeles vigtig for den fremtidige løn. Tidligere så vi, at sprog og opholdsgrundlag var vigtige risikofaktorer, men her mistede de deres betydning, efter vi tog højde for indvandrernes historiske tilknytning til arbejdsmarkedet og deres oprindelsesland. Vi fandt også, at det påvirkede arbejdsmarkedstilknnytningen negativt for indvandrere, hvis de havde oprindelse i bl.a. Irak, Somalia, Afghanistan, Libanon og Syrien. Endelig fandt vi meget stærke indikationer for, at dårligt psykisk helbred er en vigtig risikofaktor i arbejdsmarkedssammenhæng.

Den samlede model for indvandrere prædikerer korrekt i 3/4 af tilfældene, hvilket må anses som tilfredsstillende. Grunden til, at modellen forudsiger så godt, er, at vi har inkluderet arbejdsmarkedsforhold, der er meget højt korreleret med løn.

Vi har nu gennemgået, hvordan vores fulde sæt af forklarende variable mindsker eller øger sandsynligheden for at have en lønindkomst over 125.000 kr. i det første halvår af 2011. I næste afsnit undersøger vi, hvordan samme fulde sæt af forklarende variable korrelerer med sandsynligheden for ikke at have en lønindkomst i samme tidsperiode.

4.3 Udvidet model for ingen lønindkomst

Såvel population som den statistiske metode er nøjagtig den samme som ovenfor – kun definitionen af udfaldsvariablen er ændret. I forhold til den indledende analyse, som vi præsenterede i Tabel 4.1, er forskellen, at vi nu inkluderer flere kontrolvariable, der muligvis kan ændre betydningen af de demografiske, socioøkonomiske og helbredsmæssige nøglevariable.

Som ovenfor, så omfatter de ekstra kontrolvariable bl.a. historisk og nuværende arbejdsmarkedstilknytning samt oprindelseslandet for indvandrerne.

I udgangspunktet forventer vi, at koefficienterne skifter fortegn, således at en odds-ratio, der i regressionen for en lønindkomst på 125.000 kr. var større end en, nu bliver mindre end en. Eller sagt anderledes: Hvis psykisk sygdom sænker chancen for at have en lønindkomst over 125.000 kr., så vil det være naturligt at forvente, at psykisk sygdom øger risikoen for ikke at have nogen lønindkomst.

Resultaterne for de demografiske og socioøkonomiske variable viser i store træk det forventede. En højere alder øger risikoen for ikke at have en lønindkomst både for danskere, indvandrere og efterkommere (om end insignifikant), mens kvinder og ugifte har højere risiko end mænd og gifte personer. Ingen børn samt fire eller flere hjemmeboende børn, høj årsledighedsgrad og lav erhvervsgrad øger også risikoen. Det er positivt for indvandrere at bo i kommunerne umiddelbart vest for hovedstaden, mens det trækker ned at bo i de store provinskommuner i Jylland og på Fyn. Hverken indvandrerandelen eller den lokale ledighed påvirkes i nævneværdig grad.

Bedre uddannede har lavere risiko for ikke at have en lønindkomst, mens den store gruppe af indvandrere uden uddannelsesoplysninger ikke adskiller sig signifikant fra gruppen, der kun har en grundskolebaggrund. For indvandrere er mønstret nu mere tydeligt, end det vi så indledningsvist i Tabel 4.1. En højere uddannelse ser dog ud til at have en marginalt større betydning for danskere end for indvandrere. På det punkt ligner efterkommerne danskerne. Når vi ser på betydningen af stillingsstrukturen, så ser vi, at en lavere tilknytning til arbejdsmarkedet i meget høj grad prædikerer en nul-lønindkomst. Vi bemærker igen, at odds-ratioen for selvstændige er misvisende, fordi deres indkomst ikke registreres som lønindkomst.

Tabel 4.7 Logistisk regression for nul-lønindkomst i 2011; demografiske og socioøkonomiske variable

	Danskere	Indvandrere			Efterkommere
		Alle	Mænd	Kvinder	
25-29 år	0,762**	0,947*	0,805**	1,048	0,872+
30-34 år	0,828**	0,960+	0,819**	1,068*	0,938
40-44 år	1,221**	1,070**	1,172**	0,996	0,875
45-49 år	1,469**	1,206**	1,389**	1,077+	1,349
50-54 år	1,610**	1,362**	1,535**	1,209**	1,321
55-59 år	1,849**	1,554**	1,859**	1,302**	1,706
Mand	0,937**	0,779**			0,803**
Ikke gift	1,217**	1,284**	1,191**	1,350**	1,184**
Ingen hjemmeboende børn	1,258**	1,257**	1,310**	1,260**	1,164*
1 barn	1,105**	1,041+	1,019	1,083**	1,096
3 børn	1,016	1,091**	1,020	1,108**	1,211+
4 børn eller flere	1,114**	1,276**	1,267**	1,246**	1,125
Årsledighedsgrad 1-19%	1,254**	1,182**	1,181**	1,182**	1,686**
Årsledighedsgrad 20% eller derover	1,258**	1,185**	1,118**	1,248**	1,736**
Erhvervsgrad mellem 33% og 65%	0,535**	0,518**	0,477**	0,488**	0,453**
Erhvervsgrad 66% eller derover	0,281**	0,339**	0,284**	0,340**	0,212*
Har modtaget sygedagpenge	1,081**	1,122**	1,127**	1,153**	1,021
København	1,058**	1,069*	1,122*	1,021	1,203+
Vestegnskommuner	1,039**	0,895**	0,939	0,866**	1,017
Aarhus	1,028*	1,096**	1,108*	1,129**	1,126
Aalborg	1,062**	1,125**	1,191**	1,077	1,017
Odense	1,109**	1,151**	1,259**	1,071	1,215
Gymnasial uddannelse	0,875**	0,915**	0,875**	0,904**	0,852*
Erhvervsfaglig uddannelse	0,756**	0,817**	0,854**	0,770**	0,628**
Kort videregående uddannelse	0,743**	0,851**	0,790**	0,896*	0,677**
Mellemlang videregående uddannelse	0,573**	0,754**	0,787**	0,701**	0,455**
Lang videregående uddannelse	0,616**	0,838**	0,797**	0,856**	0,551**
Manglende uddannelsesoplysninger	1,172**	0,973	0,964	0,960	1,022
Topleder el. lønmodtager på højeste niv.	0,772**	0,752**	0,756**	0,682**	0,803
Øvrige lønmodtagere	0,996	0,933*	1,162**	0,766**	0,889
Selvstændig	29,328**	24,478**	24,506**	22,207**	43,119**
Ledige (uge 48)	2,734**	2,757**	3,130**	2,486**	6,061+
Syge- og barselsdagpenge	5,601**	5,020**	7,851**	4,014**	10,276*
Aktivering og rehabilitering	6,440**	4,673**	5,429**	4,347**	10,486*
Kontanthjælp og integrationsydelse	22,569**	13,007**	16,877**	11,627**	40,455**
Øvrige uden for arbejdsstyrken	8,133**	10,888**	9,764**	11,516**	17,609**
Under uddannelse	4,775**	4,191**	4,911**	3,830**	9,791*
* Indvandrerandel i kom. (kvintil)	1,017**	1,028**	1,009	1,027*	0,998
* Lokal ledighed	0,998	0,989	0,991	0,989	1,026
Antal personer	2.128.836	167.801	78.766	89.035	13.651

Baseline er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i 2009. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. Hun har 25-50%'s risiko for at opleve sproglige vanskeligheder. '+' indikerer, at odds-ratioen er signifikant forskellig fra en på et 10%'s niveau, mens '**' og '***' angiver signifikans på hhv. 5% og 1%'s niveau.

Tabel 4.8 viser resultaterne for de indvandrerspecifikke variable. Tidligere så vi, at personer, der havde opholdt sig længere tid i Danmark, havde en højere chance for at have en lønindkomst over 125.000 kr. Tabellen nedenfor viser, at personer, der har været kort tid i Danmark, har en højere risiko for ikke at have nogen lønmodtagerindkomst i første halvår af 2011. Igen er det specielt kvinder fra lande som Irak, Libanon, Pakistan og Somalia, der skil-

ler sig negativt ud, mens mænd fra Iran, Pakistan og Afghanistan skiller sig positivt ud. Kvinder fra Kina og Filippinerne skiller sig positivt ud. Både mænd og kvinder fra det tidligere Jugoslavien og de tidligere Østbloklende, Vietnam og Thailand har en markant lavere risiko for ikke at have nogen lønindkomst.

Tabel 4.8 Logistisk regression for nul-lønindkomst i 2011; indvandrerspecifikke variable

	Danskere	Indvandrere			Efterkommere
		Alle	Mænd	Kvinder	
Højst 4 år siden indrejse		0,823**	0,918*	0,768**	
10-14 år siden indrejse		0,934**	0,928*	0,944+	
15-19 år siden indrejse		0,936**	0,961	0,926*	
Min. 20 år siden indrejse		0,979	1,022	0,936	
Iran		0,894*	0,856*	0,934	1,000
Irak		1,227**	1,230**	1,269**	1,891+
Vietnam		0,700**	0,875*	0,593**	0,787
Libanon		1,284**	1,083	1,565**	1,498+
Pakistan		0,950	0,754**	1,217**	0,994
Bosnien		0,867**	0,993	0,751**	0,978
Jugoslavien		1,090+	1,155*	0,997	0,995
Makedonien		0,705**	0,640**	0,787*	0,815
Kroatien		1,039	0,930	1,126	0,966
Serbien-Montenegro		1,003	1,051	0,947	
Thailand		0,490**	0,728**	0,473**	0,536+
Somalia		1,324**	1,385**	1,344**	1,011
Sri Lanka		0,645**	0,660**	0,632**	1,505
Filippinerne		0,917*	0,985	0,817**	0,614+
Marokko		0,905+	0,884+	0,919	0,946
Afghanistan		0,928	0,849*	1,084	0,978
Kina		0,936	1,254**	0,777**	0,638
Ukraine		0,655**	0,803*	0,609**	0,584
Rumænien		0,442**	0,462**	0,450**	0,804
Indien		0,991	1,096	0,980	1,202
Rusland		0,811**	1,059	0,751**	0,318
Brasilien		0,726**	1,283+	0,601**	0,921
Bulgarien		0,389**	0,440**	0,372**	0,566
Syrien		1,204*	1,150	1,302*	0,982
Ghana		0,610**	0,800+	0,432**	0,555
Nepal		0,504**	0,585**	0,412**	2,886
Øvrige ikke-vestlige lande		0,836**	0,904*	0,788**	0,948
Flygtning eller vurderet flygtning		1,078**	0,999	1,131**	
Lav risiko for sprogproblemer		0,868**	0,947	0,778**	
Mellemhøj risiko for sprogproblemer		1,042+	0,983	1,110**	
Høj risiko for sprogproblemer		1,061	0,767**	1,185**	
Antal personer	2.128.836	167.801	78.766	89.035	13.651

Baseline er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i 2009. Hun har 25-50%’s risiko for at opleve sproglige problemer. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. '+' indikerer, at odds-ratioen er signifikant forskellig fra et på et 10%’s niveau, mens '**' og '**' angiver signifikans på hhv. 5% og 1%’s niveau.

For kvinder øger det risikoen for ikke at have nogen lønindkomst med 13%, hvis de er flygtet hertil, og det påvirker også signifikant negativt, hvis de er dårlige til dansk. For mænd er odds-ratioen for flygtningestatus og gode sprogkompetencer statistisk insignifikant og tæt på en, hvilket indikerer, at deres risiko for ikke at være i lønnet beskæftigelse ikke afhænger af, om de er gode til sprog eller flygtninge. Er mændene dårlige til sprog, har de derimod laverer risiko for ikke at have nogen løn, hvilket vi ikke har nogen intuitiv forklaring på.

Tabel 4.9 viser resultaterne for de forskellige helbredsindikatorer. Vi vil her nævne, at specielt indikatorer for psykisk sygdom slår stærkt ud i modellen. Tidligere så vi eksempelvis, at køb af psykofarmaka sænkede chancen for at tjene 125.000 kr. eller mere med omkring 25-30%. Her ser vi, at samme midler øger risikoen for at have en nul-lønindkomst med 54-63%. For mandelige indvandrere er det specielt grupperne insulin, hjerteterapi, smertestillende midler (N02), psykofarmaka (N05 og N06) og midler mod afhængighed, der udgør risikofaktorer for ikke at have nogen lønindkomst.

Tabel 4.9 Logistisk regression for nul-lønindkomst i 2011; lægemidler, diagnoser og sygesikrings sygdomme

	Danskere	Indvandrere			Efterkommere
		Alle	Mænd	Kvinder	
Antacid (A02)	1,088**	1,043*	1,018	1,057*	0,889
Insulin (A10)	1,180**	1,235**	1,241**	1,233**	1,115
Antithrombose (B01)	1,047*	0,894+	0,972	0,876	1,114
Hjerteterapi (C01)	1,047	1,336**	1,377**	1,227	0,856
Diuretica (C03)	1,054**	1,052	1,072	1,085	1,370
Betablokkere (C07)	1,068**	1,026	1,024	1,003	0,799
Calciumantagonister (C08)	1,030+	1,091	1,125	1,053	1,813
Renin-angiotensin syst. (C09)	0,997	1,006	0,955	1,073	1,042
Lipid-modif. præp. (C10)	0,981	1,022	0,987	1,073	1,099
Antireumatica og antiinflam. (M01)	1,046**	1,067**	1,047+	1,079**	1,020
Analgetica (N02)	1,221**	1,255**	1,186**	1,301**	1,256**
Pyskoleptica og psykoanaleptica (N05-6)	1,571**	1,597**	1,553**	1,627**	1,534**
Anti-afhængighed (N07B)	1,419**	1,347*	1,834**	0,677	2,504*
KOL-præparater (R03)	1,045**	0,976	0,975	0,975	1,050
Antihistaminer (R06)	0,975*	0,965	0,901*	1,015	1,069
Infektioner (A, B)	1,088**	0,999	0,966	1,029	0,992
Svulster (C00-D48)	1,049**	0,931	1,132	0,908+	0,784
Blodsygdomme mv. (D50-D89)	1,146**	1,051	1,141	1,020	0,828
Ernæringssygdomme mv. (E)	1,063**	1,047	0,909	1,109*	0,993
Psykiske lidelser (F)	1,485**	1,310**	1,450**	1,169	1,157
Nervesygdomme (G)	1,211**	1,271**	1,076	1,415**	1,174
Øre- og øjenssygdomme (H)	1,104**	1,031	1,134+	0,994	1,153
Kredsløbssygdomme (I)	1,079**	1,141**	1,085	1,176*	1,019
Åndedrætssygdomme (J)	1,097**	1,046	1,058	1,027	1,115
Fordøjelsessygdomme (K)	1,029*	1,073*	1,036	1,102*	1,039
Hudsygdomme (L)	1,090**	0,904*	1,050	0,794**	1,406**
Knoglesygdomme mv. (M)	1,156**	1,347**	1,255**	1,450**	1,124
Urinvejssygdomme mv. (N)	0,987	0,963	0,980	0,966	1,141
Svangerskabs., misdan. mv. (O, P, Q)	0,985	0,963	1,060	0,943	0,885
Div. sympt. (mavesmerte, besvim.) (R)	1,106**	1,101**	1,173**	1,051	1,185+
Ydre påvirkning (ulykke mv.) (S, T, X, Y)	1,053**	1,085**	1,074+	1,113*	1,014

	Danskere	Indvandrere			Efterkommere
		Alle	Mænd	Kvinder	
Undersøgelse (Z00-Z13)	1,058**	1,080**	1,084*	1,078*	1,057
Kontakt ifm. reproduktion (Z30-Z39)	1,129**	1,135**	1,122	1,113**	1,488**
Øvrige forhold (Z20-Z29, Z40-Z99)	1,038**	1,058	0,976	1,110*	0,861
*Alm. læge (antal kontakter)	1,006**	1,002	1,003	1,001	0,998
Tandlæge	0,773**	0,868**	0,881**	0,870**	0,928
Fysioterapi	1,009	1,029	1,055	1,033	1,147
Kiropraktik	0,921**	0,874**	0,849*	0,917	0,808
Psykologhjælp	1,207**	1,101**	1,045	1,138**	1,151
Psykiatri	1,409**	1,456**	1,307**	1,570**	1,004
Hud- og kønssygdomme	0,965**	0,994	1,010	0,972	1,116
Øjenlæge	0,977+	1,032	1,033	1,027	1,047
Ørelæge	0,975*	0,946*	0,962	0,924*	0,920
Gynækologi	0,944**	0,940+	0,962	0,938+	0,822+
Reumatologi	1,204**	1,197**	1,081	1,288**	0,902
Intern medicin	0,975	0,927	0,979	0,893	1,009
Kirurgi (plastik, ortopædisk mv.)	0,984	1,028	1,034	1,023	1,100
Antal personer	2.128.836	167.801	78.766	89.035	13.651

Baseline er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i 2009. Hun har 25-50%'s risiko for at opleve sproglige problemer. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. '+' indikerer, at odds-ratioen er signifikant forskellig fra 1 på et 10%'s niveau, mens '*' og '**' angiver signifikans på hhv. 5% og 1%'s niveau.

For indvandrere slår diagnoser for psykiske lidelser negativt ud for mændene, mens nervesygdomme virker negativt for kvinder. Diagnoser for knoglesygdomme virker negativt både for mænd og kvinder. Også når vi ser på sygesikringsydelser, slår mentale lidelser og knoglesygdomme (reumatologi) negativt ud.

Ligesom tidligere har vi også i denne model estimeret en alternativ specifikation for helbredsindikatorerne, der giver et mindre nuanceret billede. De alternative helbredsindikatorer er samme tredeling, som vi benyttede ovenfor. Også her slår psykisk sygdom stærkest ud. Personer med psykisk sygdom har således 75-86% højere risiko for ikke at have nogen lønindkomst. Vi bemærker, at risikoen, når den er defineret som her, ikke adskiller sig væsentligt for mandlige indvandrere, kvindelige indvandrere og danskere. Knogle-, muskel- og skeletsygdomme øger risikoen med 57% for indvandrerkvinder og 33% for mændene, mod 31% for danskerne. Hjerter-kar-sygdomme øger risikoen for ikke at have nogen lønindkomst med mellem 14 og 23%.

Tabel 4.10 Alternativ specifikation af helbredsindikatorer

	Danskere	Indvandrere		
		Alle	Mænd	Kvinder
Psykisk sygdom	1,859**	1,777**	1,750**	1,796**
Skelet-, knogle- og muskelsygdom	1,313**	1,452**	1,333**	1,568**
Hjerte-kar-sygdom	1,141**	1,179**	1,142**	1,226**
Antal personer	2.128.836	167.801	78.766	89.035

Baseline er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har høj erhvervsgrad og hverken modtaget sygedagpenge eller været ledig. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. '+' indikerer, at odds-ratioen er signifikant forskellig fra en på et 10%'s niveau, mens '**' og '***' angiver signifikans på hhv. 5% og 1%'s niveau.

Generelt kan vi sige, at de faktorer, der øger risikoen for at tjene under 125.000 kr., også er de faktorer, der øger risikoen for, at en person ikke har nogen lønindkomst og dermed enten modtager overførselsindkomst, lever af opsparing eller er forsørget af familien. Ældre, kvinder og ugifte har højere risiko for ikke at være lønmodtagere, mens det også øger risikoen at være lavt uddannet, have høj ledighed eller lav erhvervs erfaring. Som vi så tidligere, er der også højere risiko, hvis en indvandrer har opholdt sig kort tid i Danmark, og hvis indvandreren har oprindelse i visse mellemøstlige lande. For kvinderne trak det ned bl.a. at komme fra Irak, Libanon, Pakistan, Somalia og Syrien, mens det for mænd var positivt at komme fra fx Bulgarien, Pakistan, Thailand og Sri Lanka. Vi kunne også konstatere, at det for kvinder øgede risikoen, hvis de var flygtet hertil eller havde sprogproblemer, hvorimod det ikke havde negativ virkning for mænd. Når vi ikke kontrollerer for oprindelsesland og historisk arbejdsmarkedstilknnytning, så er sprog og flygtningestatus meget vigtige risikofaktorer. Med hensyn til helbredet, var det specielt psykisk sygdom, der slog negativt ud i vores undersøgelse, men også reumatiske sygdomme havde en betydelig negativ virkning.

I næste afsnit undersøger vi resultaternes stabilitet over tid og dermed også, om der er indikationer for, at resultaterne er afhængige af den økonomiske konjunktur.

4.4 Udvidet model for lønindkomst over 120.000 kr. i 2008

For at teste resultaternes generelle følsomhed, men også for at se om der kan spores nogen markante ændringer før og efter den økonomiske krise, har vi kørt en regression, der stort set er identisk med den ovenfor, men blot skruet tiden tre år tilbage. Det betyder, at vi ser på lønindkomsten i perioden januar til juni 2008 i stedet for 2011. Vi benytter stadigvæk medianlønningen for en ufaglært pædagogmedhjælper i en vuggestue eller børnehave som udgangspunkt, når vi definerer indikatorvariablen for en rimelig lønindkomst, men i denne periode er niveauet 5.000 kr. lavere, end vi så ovenfor.

Det første, vi vil bemærke i relation til Tabel 4.11, er, at indvandrerpopulationen her kun udgør godt 146.000 personer mod knap 168.000 to år senere. Det vil sige, at resultaterne i tabellen også er udtryk for en ændret befolkningssammensætning. Denne ændring er en konsekvens af en nettoindvandring, og når vi undersøger aldersfordelingen (ikke vist), kan vi se,

at der er kommet flere unge indvandrere, færre i alderen mellem 35 og 54 år og lidt flere personer over 50 år. Denne ændring kan i sig selv påvirke estimaterne.

Tabel 4.11 Logistisk regression for lønindkomst i 2008; demografiske og socioøkonomiske variable

	Danskere	Indvandrere			Efterkommere
		Alle	Mænd	Kvinder	
25-29 år	0,705**	0,684**	0,729**	0,647**	0,845
30-34 år	1,037**	0,955*	1,026	0,897**	1,049
40-44 år	0,832**	0,920**	0,863**	0,957	1,630+
45-49 år	0,713**	0,865**	0,786**	0,934+	0,734
50-54 år	0,668**	0,743**	0,685**	0,763**	0,794
55-59 år	0,549**	0,622**	0,594**	0,599**	1,002
Mand	1,694**	1,751**			1,700**
Ikke gift	0,976**	0,936**	0,849**	1,027	0,877*
Ingen hjemmeboende børn	0,775**	0,800**	0,782**	0,841**	0,892
1 barn	0,941**	0,940**	1,018	0,878**	0,912
3 børn	0,894**	0,913**	0,940	0,903**	0,937
4 børn eller mere	0,776**	0,789**	0,904*	0,668**	0,623+
Årsledighedsgrad 1-19%	0,712**	0,887**	0,837**	0,930*	0,905
Årsledighedsgrad 20% eller derover	0,400**	0,600**	0,599**	0,594**	0,688**
Erhvervsgrad mellem 33 og 66%	2,110**	2,530**	2,442**	2,684**	2,276**
Erhvervsgrad over 66%	5,761**	4,894**	4,538**	5,745**	4,897**
Har modtaget sygedagpenge	0,683**	0,787**	0,795**	0,765**	0,773**
København	0,874**	0,919*	0,839**	1,017	0,921
Vestegnskommuner	1,174**	1,103**	1,036	1,174**	1,046
Aarhus	0,809**	0,841**	0,829**	0,853**	0,776+
Aalborg	1,048**	1,029	0,905	1,172*	1,012
Odense	0,875**	0,845**	0,817**	0,871*	0,802
* Indvandrerandel i kom. (kvintil)	1,026**	1,043**	1,006	1,071**	0,961
* Lokal ledighed	0,974**	0,965**	0,992	0,939**	0,977
Gymnasial uddannelse	0,713**	0,951+	0,966	0,938+	0,943
Erhvervsfaglig uddannelse	1,457**	1,413**	1,334**	1,467**	1,877**
Kort videregående uddannelse	1,622**	1,390**	1,375**	1,381**	1,783**
Mellemlang. videregående uddannelse	1,726**	1,414**	1,463**	1,375**	2,685**
Lang videregående uddannelse	1,570**	1,552**	1,547**	1,527**	2,084**
Manglende uddannelse oplysninger	0,824**	0,979	0,983	0,987	0,827
Topleder eller lønmodtager på højeste niv.	1,605**	1,479**	1,433**	1,548**	1,701**
Øvrige lønmodtagere	0,784**	0,739**	0,690**	0,801**	0,924
Selvstændig	0,068**	0,055**	0,054**	0,059**	0,143*
Ledige (uge 48)	1,038	0,642*	0,658	0,573	0,812
Syge- og barselsdagpenge	0,511**	0,295**	0,198**	0,341**	0,416
Aktivering og rehabilitering	0,324**	0,292**	0,328**	0,256**	0,376
Kontanthjælp og int.ydelse	0,105**	0,092**	0,090**	0,088**	0,126*
Øvrige uden for arbejdsstyrken	0,211**	0,191**	0,203**	0,179**	0,219
Under uddannelse	0,472**	0,307**	0,254**	0,321**	0,365
Antal personer	2.180.527	146.220	69.990	76.209	10.146

Baseline er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i 2006. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. '+' indikerer, at odds-ratioen er signifikant forskellig fra en på et 10%-s niveau, mens '**' og '***' angiver signifikans på hhv. 5% og 1%-s niveau.

I forhold til de ovenstående 2011-estimer er det mest overraskende, at resultaterne for indvandrere ved første øjekast ikke adskiller sig i en statistisk set nævneværdig grad. Men der er dog mindre afvigelser.

Én undtagelse er odds-ratioen for årsledighedsgradsvariablen. Vi så tidligere, at en årsledighedsgrad under 1/5 sænkede sandsynligheden for, at indvandrere tjente over 125.000 kr. med godt 45% relativt til personer, der ikke var ledige, mens en højere årsledighedsgrad sænkede sandsynligheden med 62%. Før krisen viser resultaterne imidlertid, at førstnævnte variabel kun sænker chancen med 11%, mens den højeste årsledighed sænker chancen med 40%. En forsigtig fortolkning af, at høj årsledighed er en vigtigere determinant for lønnen i en lavkonjunktur end i en højkonjunktur, kan være, at indvandrere med høj ledighed har vanskeligere ved at opretholde lønniveauet i en lavkonjunktur end i en højkonjunktur. Desværre er det meget usikkert, om forskellen faktisk skyldes konjunktur eller en anderledes sammensætning af personer.

Vi har ikke valgt at vise estimerne for de øvrige variable, da resultaterne i det store og hele ikke adskiller sig fra dem, vi allerede har gennemgået ovenfor. Vores konklusion er derfor, at vores resultater er stabile i forhold til økonomiske konjunkturcykler og i forhold til små ændringer i befolkningssammensætningen.

Vi har nu set på, hvilke faktorer der forklarer chancen for at tjene en minimumsløn samt risikoen for ikke at have nogen lønmodtagerindkomst. Det sidste hovedresultat, vi vil gennemgå, omhandler, hvilke faktorer der påvirker risikoen for permanent at forlade arbejdsmarkedet til en førtidspension.

4.5 Model for overgang til førtidspension i 2007–2012

Vi vender os nu væk fra løn som udfaldsvariabel og kigger på de personer, som helt falder ud af arbejdsmarkedet ved at overgå til førtidspension i perioden mellem 2007 og 2012. Vores udgangspunkt er den population, vi har udtrukket i 2006, som er kendetegnet ved ikke at være på førtidspension på daværende tidspunkt. Den logistiske regressionsmodels højre side – dvs. de forklarende variable – er nøjagtig den samme, som vi netop har set ovenfor, men venstre side, eller den variabel som vi forsøger at forklare, er skiftet ud med en indikatorvariabel, der er lig med en, hvis personen overgår til førtidspension i de følgende seks år, og ellers nul. I løbet af 5 ½ år er det 3,4% af danskerne, 10,4% af indvandrere og 2,2% af efterkommerne, der tilkendes førtidspension. Sidstnævntes lave andel i kombination med en lille population gør, at estimerne overvejende er insignifikante, hvorfor vi har valgt ikke at vise dem i tabellerne nedenfor.

I udgangspunktet forventer vi, at denne regression vil vise, at ældre, personer med lavere uddannelse og lav tilknytning til arbejdsmarkedet samt individer med dårligere helbred har en højere risiko for at trække sig ud af arbejdsmarkedet permanent. Resultaterne i Tabel 4.12 viser, at risikoen for at overgå til førtidspension er stærkt stigende med alderen. 25-29-årige indvandrere har således 66% lavere sandsynlighed for at gå på førtidspension end en 35-39-

årig, mens en 55-59-årig har en næsten fire gange så høj risiko. Danske mænd har en 19% højere risiko for at overgå til førtidspension, mens indvandrermand har en 34% højere risiko.

Civilstand har en signifikant betydning for danskere, men ikke for indvandrere. Danske, der ikke er gift, har således en 39% højere risiko for at gå på førtidspension. Personer uden hjemmeboende børn og indvandrere med tre eller flere hjemmeboende børn har også højere risiko end personer med netop et eller to hjemmeboende børn.

Vi noterer, at personer med højere erhvervsgrad har mindre risiko for at overgå til førtidspension, men samtidig også at personer, der har været ledige i 2006, har mindre risiko for at overgå til førtidspension end personer, der ikke har været ledige i 2006. Sidstnævnte kan skyldes, at vejen til førtidspension ofte er langstrakt, så personer, der få år forinden har været på dagpenge, er tættere tilknyttet arbejdsmarkedet, hvorimod gruppen af personer, der ikke har været ledige, har en overrepræsentation af kontanthjælpsmodtagere og øvrige uden for arbejdsstyrken.¹⁴

Både indvandrere og danskere, der bor i hovedstadskommunerne, har en markant lavere risiko for at blive førtidspensioneret, mens borgere i de to største provinskommuner har en højere risiko. Specielt er det bemærkelsesværdigt, at indvandrere bosat i Aarhus Kommune har en 2,6 gange så høj risiko som personer bosat i den øvrige provins.

Samtidig har højere uddannede en lavere risiko for at gå på førtidspension end personer med en grundskole. Dette gælder også for den store gruppe af indvandrere med manglende uddannelsesoplysninger. For danskerne er højere uddannelse dog en vigtigere faktor for en lav risiko, end det er for indvandrerne.

Kontanthjælpsmodtagere og personer, som har været på sygedagpenge, har en 3-4 gange så høj risiko som lønmodtagere på mellem- eller grundniveau. Det er bemærkelsesværdigt, at danske lønmodtagere på øverste niveau har en 14% højere risiko for at få førtidspension, mens risikoen er højere for de indvandrede mænd, om end statistisk insignifikant. Den markant højere risiko, vi ser for kvindelige lønmodtagere på højeste niveau, kan udelukkende tilskrives nogle få såkaldte outliers.

¹⁴ De personer, der ikke har været ledige i 2006, er således en mere heterogent sammensat gruppe, der også består af personer med meget lav arbejdsmarkedstilknytning, hvor de arbejdsløse gennemsnitligt har en lavere risiko for førtidspension. I princippet burde den underliggende niveauforskel i årsledighedsgrad mellem personer, der har forskellige grader af arbejdsmarkedstilknytning, blive opfanget af kontrolvariablen for stilling. Det ser imidlertid ud til ikke at være tilfældet, hvorfor resultaterne for årsledighedsgrad skal fortolkes meget forsigtigt i denne regression.

Tabel 4.12 Logistisk regression for førtidspension; demografiske og socioøkonomiske variable

	Danskere	Indvandrere		
		Alle	Mænd	Kvinder
25-29 år	0,304**	0,327**	0,397**	0,266**
30-34 år	0,587**	0,602**	0,635**	0,573**
40-44 år	1,596**	1,566**	1,488**	1,690**
45-49 år	2,248**	2,192**	2,151**	2,299**
50-54 år	3,033**	3,168**	3,447**	3,135**
55-59 år	3,548**	3,769**	4,005**	3,908**
Mand	1,192**	1,337**		
Ikke gift	1,385**	0,991	0,980	0,995
Ingen hjemmeboende børn	1,805**	1,327**	1,152**	1,421**
1 barn	1,212**	1,095*	0,981	1,177**
3 børn	1,049+	1,160**	1,132*	1,182**
4 børn eller mere	1,074	1,202**	1,194**	1,196**
Årsledighedsgrad 1-19%	0,748**	0,819**	0,851**	0,770**
Årsledighedsgrad 20%	0,742**	0,712**	0,714**	0,699**
Erhvervsgrad mellem 1/3 og 2/3	0,532**	0,708**	0,696**	0,818**
Erhvervsgrad over 2/3	0,299**	0,535**	0,503**	0,650**
Har modtaget sygedagpenge	1,582**	1,635**	1,499**	1,720**
København	0,356**	0,270**	0,288**	0,237**
Vestegnskommuner	0,933**	0,701**	0,681**	0,704**
Aarhus	1,343**	2,600**	2,135**	3,162**
Aalborg	0,741**	0,883+	0,815+	0,946
Odense	1,059*	1,432**	1,289**	1,558**
*Indvandrerandel i kom. (kvintil)	0,930**	0,934**	0,943**	0,929**
*Lokal ledighed	1,022**	0,984	0,992	0,981
Gymnasial uddannelse	0,473**	0,852**	0,763**	0,908+
Erhvervsfaglig uddannelse	0,617**	0,850**	0,836**	0,840**
Kort videregående uddannelse	0,565**	0,709**	0,688**	0,724**
Mellemlang videregående uddannelse	0,567**	0,807**	0,740**	0,835*
Lang videregående uddannelse	0,262**	0,624**	0,566**	0,658**
Manglende uddannelsesoplysninger	1,186**	0,883**	0,844**	0,932
Lønmodtagere på højeste niveau	1,137**	1,433*	1,152	1,962**
Øvrige lønmodtagere	0,930**	1,010	1,030	1,029
Selvstændig	0,880	0,672	0,621	1,188
Ledige (uge 48)	1,226	1,872	1,672	2,889
Syge- og barseldagpenge	2,055**	2,362+	2,552+	3,093
Aktivering og rehabilitering	1,944**	2,167	2,026	3,196
Kontanthjælp og integrationsydelse	3,323**	4,731**	4,873**	6,743+
Øvrige uden for arbejdsstyrken	0,738	1,023	1,074	1,421
Under uddannelse	0,389**	0,345*	0,287+	0,543
Antal personer	2.162.529	146.064	69.215	76.737

Referencepersonen er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i 2006. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. Hun har 25-50%'s risiko for at opleve sproglige problemer.'+' indikerer, at odds-ratioen er signifikant forskellig fra en på et 10%'s niveau, mens '*' og '**' angiver signifikans på hhv. 5% og 1%'s niveau.

Tabel 4.13 Logistisk regression for førtidspension; indvandrerspecifikke variable

	Danskere	Indvandrere		
		Alle	Mænd	Kvinder
Højst 4 år siden indrejse		0,472**	0,456**	0,442**
10-14 år siden indrejse		1,976**	1,586**	2,499**
15-19 år siden indrejse		2,077**	1,600**	2,678**
Minimum 20 år siden indrejse		2,084**	1,441**	2,787**
Iran		0,815**	0,903	0,717**
Irak		1,338**	1,695**	1,019
Vietnam		0,559**	0,610**	0,518**
Libanon		1,174**	1,222*	1,138
Pakistan		0,631**	0,701**	0,586**
Bosnien		1,949**	1,728**	2,083**
Jugoslavien		1,212**	1,171+	1,195*
Makedonien		1,015	1,014	0,896
Kroatien		1,500+	1,465	1,442
Serbien-Montenegro		0,248**	0,181**	0,308**
Thailand		0,439**	0,369*	0,371**
Somalia		0,518**	0,613**	0,444**
Sri Lanka		0,699**	0,710**	0,703**
Filippinerne		0,343**	0,456*	0,281**
Marokko		0,714**	0,727*	0,736**
Afghanistan		1,186*	1,005	1,440**
Kina		0,384**	0,262**	0,446**
Ukraine		0,348**	0,445+	0,297**
Rumænien		0,860	0,963	0,786
Indien		0,768+	0,731	0,734
Rusland		0,548**	0,978	0,424**
Brasilien		0,488**	0,163+	0,499**
Bulgarien		0,472*	0,169+	0,566
Syrien		1,278*	1,714**	1,084
Ghana		0,362**	0,272**	0,506*
Nepal		0,199	0,356	
Øvrige ikke-vestlige lande		0,768**	0,905	0,662**
Flygtning eller vurderet flygtning		1,570**	1,577**	1,556**
Lav risiko for sprogproblemer		0,875**	0,985	0,890
Mellemhøj risiko for sprogproblemer		1,034	0,909+	1,033
Høj risiko for sprogproblemer		0,922	0,791**	0,977
Antal personer	2.162.529	146.064	69.215	76.737

Referencepersonen er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i 2006. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. Hun har 25-50%'s risiko for at opleve sproglige problemer. '+' indikerer, at odds-ratioen er signifikant forskellig fra en på et 10%'s niveau, mens '*' og '**' angiver signifikans på hhv. 5% og 1%'s niveau.

Tabel 4.13 viser, hvordan de indvandrerspecifikke variable varierer med sandsynligheden for at overgå til førtidspension. Her ser vi, at personer, specielt kvinder, der har boet en årrække i Danmark, har en højere risiko for at få førtidspension, men også, at risikoen er større for mænd med oprindelse i Irak, Libanon, Bosnien og Syrien, mens den er signifikant lavere for mænd fra bl.a. Vietnam, Pakistan og Somalia. Flygtede personer har en 57% højere risiko for at overgå til førtidspension end personer, som ikke er flygtet, hvilket kan hænge sammen

med de oplevelser, som de har været udsat for. Overordnet har personer med bedre sprogkompetencer (dvs. lav risiko for sprogproblemer) også en lavere sandsynlighed for at overgå til førtidspension, mens estimatet for mænd, der har stor risiko for at opleve sprogproblemer, er det modsatte af, hvad vi ville have forventet. Vi har ikke nogen god forklaring på dette omvendt intuitive forhold.

Tabel 4.14 viser – ikke overraskende – at personer, som lider af psykisk sygdom, har en markant højere risiko for at overgå til førtidspension. For danskere er risikoen 3,4 gange så høj, hvis personen køber psykofarmaka, har været indlagt og fået en psykiatrisk diagnose eller har besøgt en psykiater eller psykolog. For indvandrere er den tilsvarende risiko en anelse højere, men ikke væsentlig forskellig fra danskernes. Både danskere og indvandrere, der lider af knogle-, muskel eller skeletsygdomme har en knap dobbelt så høj risiko, som personer uden disse lidelser. Personer med hjerte-kar-sygdomme har en ca. 1½ gang højere risiko for at overgå til førtidspension.

Tabel 4.14 Alternativ specifikation af helbredsindikatorer

	Danskere	Indvandrere		
		Alle	Mænd	Kvinder
Psykisk sygdom	3,400**	3,634**	3,990**	3,297**
Knogle-, muskel- eller skeletsygdom	1,734**	1,945**	1,894**	1,944**
Hjerte-kar-sygdom	1,616**	1,455**	1,443**	1,462**
Antal personer	2.162.529	146.064	69.215	76.737

Referencepersonen er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i 2006. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. Hun har 25-50%'s risiko for at opleve sproglige problemer. '+' indikerer, at odds-ratioen er signifikant forskellig fra en på et 10%'s niveau, mens '*' og '**' angiver signifikans på hhv. 5% og 1%'s niveau.

Tabel 4.15 viser et mere detaljeret billede, end det vi så i Tabel 4.14. Her kan man således se, hvor meget de enkelte medicinprodukter, diagnoser og speciallægebesøg øger (eller sænker) risikoen for at overgå til førtidspension. Overordnet understøtter tallene det billede, vi lige har tegnet af, at psykisk sygdom, hjerte-kar- og knoglesygdomme er store risikofaktorer. Til denne liste kan man tilføje nervesygdomme, infektioner samt øjen- og øresygdomme. For besøg hos ydere under den offentlige sygesikring ses, at tandlægekonsultationer og besøg hos kiropraktorer sænker risikoen for førtidspension, men da disse ydere har en høj grad af egenbetaling, så ser vi det som udtryk for selektion af borgere, som er tættere på arbejdsmarkedet og har en højere indkomst.

Tabel 4.15 Logistisk regression for førtidspension; helbredsindikatorer

	Danskere	Indvandrere		
		Alle	Mænd	Kvinder
Antacid (A02)	1,153**	1,037	1,006	1,061
Insulin (A10)	1,312**	1,101+	1,038	1,166+
Antithrombose (B01)	1,176**	0,995	1,152	0,877
Hjerteterapi (C01)	1,158**	1,041	1,068	0,976
Diuretica (C03)	1,208**	1,217**	1,429**	1,138*
Betablokkere (C07)	1,007	0,897*	0,804**	0,956
Calciumantagonister (C08)	1,129**	1,176*	1,076	1,263*
Renin-angiotensin syst. (C09)	1,040*	1,038	1,045	1,003
Lipid-modif. præp. (C10)	0,963+	1,043	1,004	1,070
Antireumatica og antiinflamm. (M01)	1,046**	1,110**	1,053	1,153**
Analgetica (N02)	1,620**	1,499**	1,661**	1,371**
Pyskoleptica og psykoanaleptica (N05-6)	1,971**	2,050**	2,054**	1,971**
Anti-afhængighed (N07B)	1,329**	0,970	0,869	1,282
KOL-præparater (R03)	1,157**	0,999	0,988	1,000
Antihistaminer (R06)	1,019	0,953	0,945	0,963
Infektioner (A, B)	1,295**	1,257**	1,247*	1,253*
Svulster (C00-D48)	1,531**	1,083	1,124	1,060
Blodsygdomme mv. (D50-D89)	1,569**	1,221+	0,979	1,345*
Ernæringssygdomme mv. (E)	1,146**	1,162**	1,252*	1,127+
Psykiske lidelser (F)	1,490**	1,362**	1,488**	1,269*
Nervesygdomme (G)	1,750**	1,405**	1,511**	1,359**
Øre- og øjenssygdomme (H)	1,449**	1,284**	1,250**	1,277**
Kredsløbssygdomme (I)	1,164**	1,180**	1,174*	1,180*
Åndedrætssygdomme (J)	1,317**	1,098+	1,133	1,052
Fordøjelsessygdomme (K)	1,120**	1,056	1,033	1,064
Hudsygdomme (L)	1,203**	1,253**	1,095	1,367**
Knoglesygdomme mv. (M)	1,237**	1,546**	1,466**	1,606**
Urinvejssygdomme mv. (N)	1,012	0,953	1,023	0,933
Svangerskabs., misdannelser mv. (O, P, Q)	0,893**	0,734**	1,352	0,793**
Div. sympt. (mavesmerter, besvimelse mv.) (R)	1,200**	1,100**	1,072	1,127**
Ydre påvirkning (ulykke mv.) (S, T, X, Y)	1,048**	1,143**	1,083	1,187**
Undersøgelse (Z00-Z13)	1,204**	1,149**	1,150**	1,147**
Kontakt ifm. reproduktion (Z30-Z39)	0,807**	0,790**	1,112	0,833**
Øvrige forhold (Z20-Z29, Z40-Z99)	1,274**	1,164**	1,206*	1,135+
Organiske psykiske lidelser (F00-09)	4,080**	3,886**	3,417*	4,308*
Brug af psyk.akt. stoffer (F10-19)	1,040	1,307+	1,150	1,669
Skizofreni, psykoser etc. (F20-29)	8,945**	6,896**	7,082**	6,095**
Affektive lidelser (F30-39)	1,744**	1,306**	1,122	1,521**
Stressrelaterede lidelser (F40-49)	1,041	1,405**	1,732**	1,201*
Adfærdssændringer (fys. bet.) (F50-59)	1,341**	1,351	0,694	2,379+
Personlighedsforstyrrelser (F60-69)	1,795**	1,799**	1,732**	1,793*
Mental retardering (F70-79)	10,699**	6,673*	5,464	8,589*
Udviklingsforstyrrelser (F80-89)	9,670**	39,316**	38,526**	
Sygdom pga. forhold i barndom (F90-98)	2,426**	0,994	0,762	1,968
Uspecificerede psykiske lidelser (F99)	1,175*	1,200	1,312	1,091
*Alm. læge (antal kontakter)	1,031**	1,026**	1,038**	1,022**
Tandlæge	0,824**	0,895**	0,856**	0,918*
Fysioterapi	1,370**	1,225**	1,221**	1,223**
Kiropraktik	0,792**	0,847*	0,740**	0,932
Psykologhjælp	1,275**	1,338**	1,514**	1,250**

	Danskere	Indvandrere		
		Alle	Mænd	Kvinder
Psykiatri	2,043**	2,172**	1,963**	2,396**
Hud- og kønssygdomme	0,967+	0,956	0,953	0,960
Øjenlæge	1,072**	1,137**	1,063	1,191**
Ørelæge	0,984	1,070+	1,192**	0,990
Gynækologi	0,799**	0,927	0,710	0,971
Reumatologi	1,335**	1,384**	1,306**	1,417**
Intern medicin	0,912+	1,037	1,143	0,915
Kirurgi (plastik, ortopædisk mv.)	0,846**	0,918+	0,806**	0,990
Antal personer	2.162.529	146.064	69.215	76.737

Referencepersonen er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager på mellem- eller grundniveau, har lav erhvervsgrad, men har hverken modtaget sygedagpenge eller været ledig i 2006. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. Hun har 25-50%'s risiko for at opleve sproglige problemer. '+' indikerer at odds-ratioen er signifikant forskellig fra en på et 10%'s niveau, mens '*' og '**' angiver signifikans på hhv. 5% og 1%'s niveau.

Opsamlende kan vi sige, at resultaterne for, hvilke faktorer der forklarer overgangen til førtidspension, på nogle områder bekræfter det billede, vi har set i de tidligere analyser. Jo højere en persons alder er, des større er risikoen for at overgå til førtidspension. Mænd, specielt med indvandrerbaggrund, har en højere risiko for at få førtidspension, mens de havde en lavere risiko for ikke at have nogen lønindkomst. Højere uddannelse og større erhvervsgrad sænker risikoen for førtidspension. Indvandrere, der har opholdt sig i Danmark i mange år, har en langt højere risiko for at få førtidspension, hvorimod vi tidligere så, at indvandrere med længere ophold har lavere risiko for ikke at have nogen lønindkomst. Mænd fra Irak, Bosnien og Syrien havde højere risiko, mens mænd fra fx Vietnam og Thailand havde lavere risiko. For kvinderne var det indvandrere fra fx Iran, Vietnam og Pakistan, der skilte sig positivt ud, mens kvinder fra Bosnien og Afghanistan skilte sig negativt ud. Flygtninge havde højere risiko for førtidspension, mens personer uden sprogproblemer havde lavere risiko. Endelig så vi, at den absolut største risikofaktor, som forventet, var psykisk sygdom efterfulgt af reumatiske sygdomme.

Litteratur

- Andersen, H.L. (2010): *Does unemployment really kill?* Working paper, København: AKF.
- Ankestyrelsen (2010): *Førtidspensioner: Årsstatistik 2009*. København: Ankestyrelsen.
- Arendt, J.N.; C.P. Nielsen & V. Jakobsen (2012): *The relationship between pre- and post-migration qualifications and their impact on employment status*. Working paper, København: AKF.
- Browning, M.; A.M. Danø & E. Heinesen (2006): Job displacement and stress-related health outcomes. *Health Economics*, 15:1061-75.
- Caliendo, M. & S. Kopeinig (2005): *Some Practical Guidance for the Implementation of Propensity Score Matching*. IZA discussion paper no. 1588. Bonn: IZA.
- Dahl, K.M. & V. Jakobsen (2005): *Køn, etnicitet og barrierer for integration. Fokus på uddannelse, arbejde og foreningsliv*. København: Rapport 05:01. SFI.
- Danmarks Statistik (2011): *Indvandrere i Danmark, 2011*. København: Danmarks Statistik.
- Fazel, M.; J. Wheeler & J. Danesh (2005): Prevalence of serious mental disorder in 7000 refugees resettled in western countries: a systematic review. *Lancet*, 356(9467):1309-14.
- Goul Andersen, J. (2008): *Holdninger til uddannelse og arbejde blandt unge indvandrere, danskere og deres forældre*. Odense: Syddansk Universitetsforlag.
- Heckman & Borjas (1980): Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence. *Economica*, 47:247-83.
- Jahoda, M. (1981): Work, Employment, and Unemployment. *American Psychologist*, 36(2):184-91.
- Leuven, E. & B. Sianesi (2003): *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*. Se <http://ideas.repec.org/c/boc/bocode/s432001.html>
- Ministeriet for Flygtninge, Indvandrere og Integration (2011): *Medborgerskab i Danmark*. Rapport fra regeringens arbejdsgruppe for bedre integration. København: Social- og Integrationsministeriet.
- Social- og Integrationsministeriet (2011): *Fakta om integration. Status og udvikling*. København: Social- og Integrationsministeriet.
- Wooldridge (2009): *Introductory Econometrics – A Modern Approach*. Mason: Thomson Higher Education.

Bilag 1: Konstruktion af sprogindeks

Som nævnt er vores indikator for, om indvandreren oplever sprogproblemer, ikke udtryk for faktiske tilkendegivelser fra hver person, men konstrueret på baggrund af interview med et udsnit af indvandrerbefolkningen foretaget i 2009-2010. Datagrundlaget er beskrevet i rapporten "Medborgerskab i Danmark, 2011". Data omfatter 4.633 spørgeskemabesvarelser fra ikke-vestlige indvandrere. Heraf kan de 4.613 matches til personer i vores registerdatabase for befolkningen i 2009. Vi sletter de 20 personer fra spørgeskemaundersøgelsen, der ikke kan findes i befolkningsregisteret. De 4.613 personer er i alderen 18 til 97 år. Vi beholder 3.413 observationer, der i 2011 vil være i alderen 25-59 år. Af disse 3.413 personer er 659 administrativt registreret som danskere, mens 318 er registreret som efterkommere. Vi beholder kun de 2.436 personer, der er registreret som indvandrere. Af disse er 328 på førtidspension. Førtidspensionisterne slettes, og vi beholder de resterende 2.108 ikke-vestlige indvandrere til analysen.

Vi bruger fem spørgsmål fra spørgeskemaundersøgelsen. De fem spørgsmål er:

- Har du problemer med at forstå det danske sprog:
 - a. når du får breve fra kommunen?
 - b. ved nyhedsudsendelser?
 - c. i forbindelse med arbejde eller uddannelse?
- Sprog for interviewet (dansk, modersmål)
- Udfyldt af interviewer: Havde interviewede problemer med at forstå/følge spørgsmålene?

På baggrund af disse besvarelser konstruerer vi en samlet score på baggrund af besvarelsene. Scoren er 0, hvis personen ikke har nogen sprogproblemer (35%), 1 ved små sprogproblemer (20%), 2 ved nogle sprogproblemer (19%), 3 ved en del sprogproblemer (21%) og 4 ved store sprogproblemer (5%). Derpå defineres en indikatorvariabel, som indtager værdien en, hvis scoren er lig 3 eller 4, og ellers nul. Denne score er konstrueret på baggrund af rapporten fra Ministeriet for Flygtninge, Indvandrere og Integration (2011).

På baggrund af vores sæt af forklarende variable, som ligger i 2009-registerdatabasen, og som vi bruger i alle øvrige regressioner, regresserer vi nu de fulde sæt af forklarende variable på den definerede indikatorvariabel, som indtager værdien en, hvis personen har en del eller store sprogproblemer. Regressionsmetoden er en logistisk regression ligesom i de øvrige analyser. Vi beholder de variable, der er præcist estimeret, dvs. hvis koefficienten er signifikant forskellig fra nul på et 10%'s niveau. Efter at have slettet insignifikante forklaringsfaktorer køres modellen igen. Resultaterne fra denne kørsel er gengivet i Tabel B1.1.

Resultaterne i Tabel B1.1 viser, at mænd har lavere sandsynlighed for at opleve sprogproblemer, ældre har større risiko, ugifte har mindre risiko, personer med mange hjemmeboende børn har højere risiko, personer med højere erhvervsgrad har langt lavere risiko, personer med marginal tilknytning til arbejdsmarkedet – fx kontanthjælpsmodtagere – har højere

risiko, højere uddannelse og længere tid i Danmark sænker risikoen, mens stort set alle har lavere risiko for at have sprogproblemer end tyrkere.

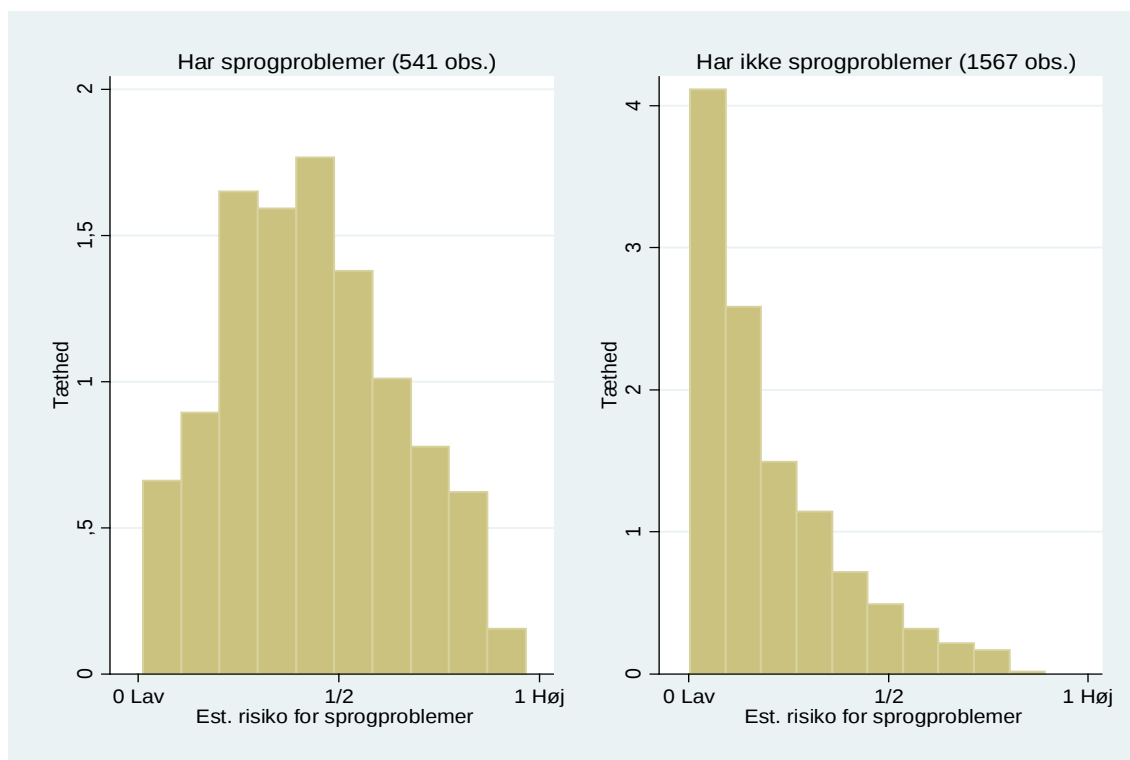
På baggrund af disse estimater beregnes nu sandsynligheden for, at hver af de 2.108 indvandrere oplever sprogproblemer. For de indvandrere, der faktisk har sprogproblemer, forventer vi, at denne sandsynlighed ligger tæt på en, mens sandsynligheden for de indvandrere, der ikke oplever nogen sprogproblemer, helst skal ligge tæt på nul. Figur B1.1 viser dette grafisk. Fordelingen af sandsynligheder for personer, der faktisk har sprogproblemer (panelet i venstre side), er nærmest normalfordelt. Hvis en estimeret sandsynlighed over $\frac{1}{2}$ karakteriseres som havende sprogproblemer, så betyder det, at modellen prædikerer korrekt i lidt under halvdelen af tilfældene. Panelet i højre side af Figur B1.1 viser fordelingen af de estimerede sandsynligheder for de personer, der ikke oplever sprogproblemer. Denne fordeling har langt den største masse i højre side, hvilket vil sige, at vi estimerer korrekt i langt de fleste tilfælde for denne gruppe.

Næste skridt er at bruge de estimerede koefficienter på hele indvandrerpopulationen, som består af 172.993 personer. For hver person beregner vi nu sandsynligheden for, at vedkommende oplever sprogproblemer, nøjagtigt på samme måde som netop beskrevet. Fordelingen af disse sandsynligheder er gengivet i Figur B1.2. Figur B1.2 adskiller sig fra Figur B1.1 ved, at der ikke opdeles på, om personen faktisk oplever sprogproblemer eller ej. Det skyldes naturligvis, at vi ikke har denne oplysning (i så fald havde hele denne øvelse nemlig været overflødig). Venstre panel i Figur B1.2 viser fordelingen blandt de 2.108 personer, hvor vi faktisk har oplysninger om sprogproblemer, mens højre panel viser fordelingen af sandsynligheder for alle 172.993 indvandrere, hvoraf vi altså kun har faktiske oplysninger for godt to tusinde borgere.

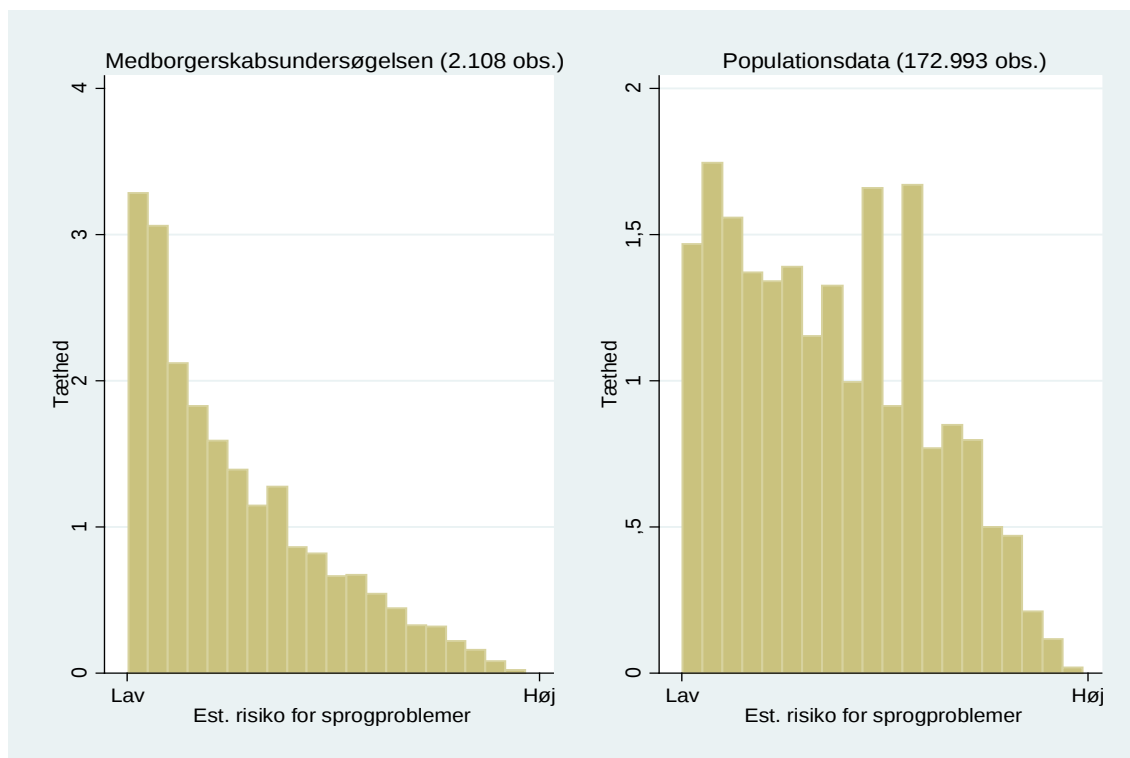
For at undgå linearitetsproblemer konstrueres nu fire indikatorvariable, som vi bruger i den endelige analyse. Tilsammen angiver de fire indikatorer, hvilket kvartil hver enkelt indvander ligger i i sandsynlighedsfordelingen. Variablen 'Q1' vil således være lig med en, hvis en person ligger i første kvartil af sandsynlighedsfordelingen, mens 'Q2' er lig med en, hvis personen ligger i andet kvartil osv. Hver indikatorvariabel er således lig med en for nøjagtigt 25% af indvandrerbefolkningen pr. konstruktion.

Vi viser i Bilag 3: Alternative modelspecifikationer, at de konstruerede sprogvariable er både økonomisk og statistisk signifikante i indkomstligningen, når vi ikke samtidig kontrollerer for oprindelsesland. Når vi inkluderer kontrolvariable for indvandrerens oprindelsesland, mister sprogvariablene deres forklaringskraft.

Figur B1.1 Fordeling af estimeret sprogrisiko på personer med og uden sprogproblemer



Figur B1.2 Fordeling af estimeret risiko for sprogproblemer i medborgerskabsundersøgelsen og populationen



Tabel B1.1 Estimation af sprogindeks

	Indvandrere fra medborgerskabsundersøgelsen
Mand	0,691**
30-34 år	1,047
40-44 år	1,351+
45-49 år	2,032**
50-54 år	1,718**
55-59 år	2,616**
Ikke gift	0,611**
4 børn eller mere	1,454+
Erhvervsgrad mellem 1/3 og 2/3	0,445**
Erhvervsgrad over 2/3	0,299**
Topleder eller lønmodtagere på højeste niveau	1,006
Øvrige lønmodtagere	1,473*
Selvstændig	1,463
Ledige (uge 48)	1,209
Syge- og barselsdagpenge	1,131
Aktivering og rehabilitering	2,438**
Kontanthjælp og int.ydelse	3,556**
Øvrige uden for arbejdsstyrken	1,691*
Under uddannelse	0,628
Gymnasial uddannelse	0,759
Erhvervsfaglig uddannelse	0,667**
Kort videregående uddannelse	0,504**
Mellemlang videregående uddannelse	0,511**
Lang videregående uddannelse	0,670
Højt specialiseret arbejde	0,176**
Arbejde, der forudsætter viden på mellemniveau	0,450**
Kontorarbejde	0,177**
Detailsalg, service- og omsorgsarbejde	0,656*
Højest 4 år siden indrejse	2,177**
10-14 år siden indrejse	1,427*
15-19 år siden indrejse	1,282
Min. 20 år siden indrejse	0,676*
Iran	0,431**
Irak	0,942
Vietnam	1,127
Libanon	0,400**
Pakistan	0,882
Bosnien	0,264**
Jugoslavien	0,271**
Makedonien	0,068*
Kroatien	0,509
Serbien-Montenegro	0,470
Antithrombose (B01)	2,742**
Anti-afhængighed (N07B)	6,498*
Nervesygdomme (G)	0,487*
Øre- og øjenssygdomme (H)	1,998*
Kredsløbssygdomme (I)	0,482*
Øvrige forhold (Z20-Z29, Z40-Z99)	0,502*
Gynækologi	1,769*
Antal observationer	2.108

Regressionens pseudo-R² er 0,213. '+' indikerer, at odds-ratioen er signifikant forskellig fra 1 på et 10%'s niveau, mens '**' og '***' angiver signifikans på hhv. 5% og 1%'s niveau.

Bilag 2: Separat analyse for selvstændiges arbejdsmarkedstilknytning

I alle rapportens analyser har vi valgt at have en samlet fremstilling for hele populationen af potentielt arbejdsdygtige danskere, indvandrere og efterkommere. Men vi har også bemærket, at fx lønindtægt må anses som et uhensigtsmæssigt mål for selvstændiges personlige indkomst, fordi løn ikke omfatter overskud ved virksomhedsdrift.

Vores datasæt indeholder imidlertid oplysninger om personers nettooverskud ved drift af selvstændig virksomhed for året 2009. Vi har derfor udtrukket alle personer, der i november 2006 var registrerede som selvstændige, og undersøgt, hvordan personlige karakteristika varierer med sandsynligheden for, at virksomheden har et nettodriftsoverskud af en vis størrelse. Grænseværdien for størrelsen af nettooverskuddet er fastsat som medianværdien, der i 2009 var 125.000 kr. Dette mål kan og skal ikke sammenlignes med den personlige lønmodtagerindkomst, som vi tidligere har behandlet. Undersøgelsen i dette bilag skal derimod ses som en følsomhedsanalyse, der har til formål at vurdere, om resultaterne for selvstændige i store træk er væsentligt anderledes end det, vi har set ovenfor.

Da vi nu kun ser på selvstændige, så må datagrundlaget nødvendigvis være mindre. I 2006 var der således godt 118.000 danskere og knap 9.000 indvandrere, som stadig var i landet i 2009. Hovedparten af de selvstændige indvandrere er mænd, og derfor viser vi ikke selvstændige regressionsestimater for kvinderne. Når populationen er mindre, betyder det specielt for indvandrerne, at vi må forvente højere standardfejl på estimerne, hvorfor der vil være flere estimer, der er statistisk insignifikante og upræcist estimeret. Nedenfor gennemgår vi kortfattet resultaterne fra følsomhedsanalysen.

For danskere viser resultaterne i Tabel B2.1 en tydelig aldersgradient, sådan at ældre har højere sandsynlighed for at drive en virksomhed, der genererer et nettooverskud på 125.000 kr. eller mere. For indvandrere ser tendensen ud til at være den samme, om end den er langt mindre udpræget. Igen ser vi, at mænd til forskel fra kvinder har en højere sandsynlighed for at drive virksomhed med et sådant overskud, mens resultaterne for ugifte indvandrere ikke er signifikant forskellig fra referencegruppen. Danskere med fire eller flere hjemmeboende børn, har en lavere chance for at have et overskud på 125.000 kr.; det samme gælder for indvandrere, men resultatet er ikke statistisk signifikant.

Uddannelseslængden har ikke nogen klar sammenhæng med virksomhedens overskud for indvandrere, men for danskere ser vi samme positive sammenhæng mellem uddannelse og overskud som tidligere.

Psykisk sygdom er derimod igen en klar risikofaktor for både indvandrere og danskere, mens knoglesygdomme og hjerte-kar-sygdomme kun har en tydelig negativ indvirkning for danskere.

Selvstændige indvandrere, der har opholdt sig længere tid i Danmark, har klart en højere sandsynlighed for at drive en virksomhed med et nettooverskud på 125.000 kr. eller mere. Statistisk set gælder det også for personer, der er indvandret fra det tidligere Jugoslavien,

Iran og Marokko sammenlignet med tyrkere. Det samme ser ud til at gælde for indvandrere fra flere andre lande, men disse resultater er ikke statistisk signifikant forskellige fra en. Kun indvandrere fra Thailand har en signifikant lavere sandsynlighed for at drive en virksomhed med et nettooverskud på 125.000 kr. eller mere sammenlignet med tyrkere. Hvis indvandre- ren er flygtet hertil, så har han en ca. 30% lavere sandsynlighed for at drive en virksomhed, der har et overskud på mindst 125.000 kr. Resultaterne for sprogvariablene er ikke statistisk signifikante, men de indikerer den samme tendens til, at selvstændige med bedre sprogkom- petencer også klarer sig bedre end selvstændige med dårligere selvvaluerede sprogkompe- tencer.

Overordnet mener vi, at denne følsomhedsanalyse for selvstændige i høj grad understøt- ter de tendenser, som vi tidligere har beskrevet.

Tabel B2.1 Resultater for selvstændige (2006). Nettooverskud ved drift af selvstændig virksomhed (2009)

	Danskere	Indvandrere	
		Alle	Kun mænd
25-29 år	0,537**	0,714**	0,690**
30-34 år	0,841**	0,843*	0,918
40-44 år	1,128**	1,111	1,012
45-49 år	1,292**	1,164+	1,064
50-54 år	1,325**	1,229+	1,091
55-59 år	1,190**	1,521**	1,366*
Mand	1,754**	1,742**	
Ikke gift	0,912**	1,029	0,986
Ingen hjemmeboende børn	0,829**	0,938	0,94
1 barn	0,941**	1,024	1,018
3 børn	0,97	1,064	1,089
4 børn eller flere	0,872**	0,873	0,848
Gymnasial uddannelse	0,997	0,972	1,004
Erhvervsfaglig uddannelse	1,556**	1,003	0,93
Kort videregående uddannelse	1,257**	0,928	1,014
Mellemlang videregående uddannelse	1,275**	0,922	0,99
Lang videregående uddannelse	2,547**	1,239	1,096
Manglende uddannelsesoplysninger	0,875+	0,972	0,955
Psyriske sygdomme	0,638**	0,672**	0,715**
Muskel-, skelet- eller knoglesygdomme	0,845**	0,987	0,998
Hjerte-kar-sygdomme	0,917**	1,039	1,008
Højest 4 år siden indrejse		1,101	0,938
10-14 år siden indrejse		1,136	1,087
15-19 år siden indrejse		1,322**	1,265*
Min. 20 år siden indrejse		1,312**	1,463**
Iran		1,379*	1,437*
Irak		1,147	1,226
Vietnam		1,340*	1,231
Libanon		1,126	1,247
Pakistan		1,676**	1,887**
Bosnien		2,248**	2,537**
Jugoslavien		1,589**	1,924**
Makedonien		2,358**	3,071**
Kroatien		1,793	3,369

	Danskere	Indvandrere	
		Alle	Kun mænd
Thailand		0,577**	0,549
Somalia		1,524	1,351
Filippinerne		1,569	1,661
Marokko		1,421*	1,635**
Afghanistan		1,319	1,266
Flygtning eller vurderet flygtning		0,729**	0,697**
Lav risiko for sprogproblemer		1,158+	0,986
Mellemhøj risiko for sprogproblemer		0,942	1,02
Høj risiko for sprogproblemer		0,838	1,14
Antal observationer	118.415	8.757	6.585

Referencepersonen er en gift, 35-39-årig kvinde med to børn. Hvis hun er indvandrer, så er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. Hun har 25-50%'s risiko for at opleve sprogvanskeligheder. Regressionens fit er hhv. 0,038, 0,041 og 0,032 målt ved pseudo-R². I tillæg til de nævnte variable, kontrollerer vi også for, om personen bor i en storbykommune og for den lokale ledighed samt indvandrerandel. Der kontrolleres for det fulde sæt af oprindelseslande, men vi har ikke inkluderet de estimer, der var mest upræcist estimeret.

Bilag 3: Alternative modelspecifikationer

Tabel B3.1 viser forskellige variationer over den model, vi indledningsvis præsenterede i Tabel 4.2 til Tabel 4.6. Modellens venstreside er en indikator for, om lønindkomsten i 1. halvår 2011 overstiger 125.000 kr.

De variable, vi har inkluderet, risikerer i nogen grad at forklare det samme, fordi de er korrelerede. Eksempelvis vil problemet opstå, hvis man inkluderer en kontrolvariabel for både, om personen er gift, og om hun er samboende. I vores tilfælde har vi flere variable, der kan være udtryk for arbejdsmarkedstilknytningen og erhvervserfaring i udgangspunktet, ligesom vi har flere variable, der kan være udtryk for kulturelle forskelle. Problemet er, at variablene påvirker hinanden i en ukendt grad, og specielt variable med stor forklaringskraft kan gøre en forskel for, om fortegnet bliver plus eller minus.

For at adressere problemet begynder vi med at estimere en model med et minimum af forklarende variable. Denne reducerede model indeholder kun de vigtigste forklarende variable. Derpå kører vi regressionen igen og tilføjer gradvist flere blokke af variable, indtil vi når til vores endelige specifikation. På den måde bliver det tydeligt, hvor meget de vigtige variable i den mest reducerede model påvirkes af, at der gradvist tilføjes flere forklarende variable.

I Tabel B3.1, der kun viser resultater for ikke-vestlige indvandrere, ser vi i første kolonne ("Alle" længst til venstre) en model, som kun kontrollerer for alder, køn, civilstand, børn, uddannelse, lokal indvandretæthed og ledighed, antal år i Danmark, opholdsgrundlag, estimerede sprogkompetencer og køb af receptpligtig medicin. Kolonne to og tre viser samme modelresultater blot opdelt efter køn. I kolonne fire ('B') udvides modellen til også at kontrollere for årsledighedsgrad, erhvervsgrad, modtagelse af sygedagpenge og en kontrolvariabel for udvalgte bopælskommuner. Sidste kolonne ('C') viser den endelige model, der tilføjer kontrolvariable for sektor, hovedbeskæftigelse, oprindelsesland, hospitalsdiagnoser og sygesikringskontakter.

Konkret kan vi se, at aldersestimaterne for 50-54-årige og 55-59-årige skifter fra at være større end en til at være mindre end en, når vi går fra model 'A' til model 'B'. Odds-ratioen for 40-44-årige og 45-49-årige falder, så de kommer tættere på en. Disse resultater betyder, at det ikke er en høj alder, der er en vigtig determinant for evnen til at opretholde en lønindkomst, men i højere grad erhvervserfaringen. Hvis vi kun havde estimeret model 'A', så havde vi konkluderet, at 55-59-årige har en 72% højere chance end en 35-39-årig for at have en indkomst over 125.000 kr., mens vi i model 'C' vil konkludere, at en 55-59-årig indvandrer har en 34% lavere chance for at have denne indkomst. Og langt hovedparten af denne ændring skyldes udelukkende inklusionen af erhvervsgraden, som udtrykker personens erhvervserfaring.

På samme måde ser vi i specifikation 'A', at flygtninge har en 26% lavere chance for at have en lønindkomst over 125.000 kr., mens personer med gode sprogkompetencer har en 2,9 gange højere chance for at have beskæftigelse til denne løn eller mere. Forklaringskraften af disse variable forsvinder imidlertid fuldstændigt, når vi i specifikation 'C' inkluderer sættet af indikatorvariable, som kontrollerer for oprindelsesland. For medicinvariablene bliver kor-

relationen mellem eksempelvis køb af psykoleptica og psykoanaleptica mindre, efter at vi inkluderer de øvrige helbredscontrollere.

Tabel B3.1 Alternative modelspecifikationer (lønindkomst over 125.000 kr. i 1. halvår 2011)

	Ikke-vestlige indvandrere				
	Reduceret model A			B: A + ledighed, erhvervsgruppe, sygedagpenge, kom.	C: Endelig specifikation
	Alle	Mænd	Kvinder		
25-29 år	0,600**	0,715**	0,482**	0,631**	0,643**
30-34 år	0,909**	1,053+	0,798**	0,957*	0,951*
40-44 år	1,342**	1,141**	1,544**	1,154**	1,021
45-49 år	1,785**	1,459**	2,092**	1,278**	0,893**
50-54 år	1,487**	1,215**	1,698**	0,993	0,807**
55-59 år	1,722**	1,483**	1,773**	0,954	0,659**
Mand	1,184**			1,203**	1,679**
Ikke gift	0,594**	0,516**	0,667**	0,656**	0,922**
Ingen hjemmeboende børn	0,879**	0,743**	1,098**	0,889**	0,770**
1 barn	1,005	1,024	0,984	1,025	0,920**
3 børn	0,808**	0,869**	0,761**	0,813**	0,924**
4 børn eller flere	0,750**	0,888**	0,611**	0,739**	0,790**
Gymnasial uddannelse	0,747**	0,820**	0,695**	0,804**	0,946*
Erhvervsfaglig uddannelse	1,112**	0,959+	1,254**	1,211**	1,452**
Kort videregående uddannelse	0,931**	0,967	0,845**	1,067*	1,332**
Mellemlang videregående uddannelse	1,241**	1,161**	1,316**	1,499**	1,579**
Lang videregående uddannelse	1,319**	1,334**	1,300**	1,626**	1,445**
Manglende uddannelsesoplysninger	0,979	1,043	0,875**	0,930**	0,997
Årslighedsgrad 1-19%				0,593**	0,545**
Årslighedsgrad 20% og derover				0,291**	0,381**
Erhvervsgrad mellem 1/3 og 2/3				3,195**	2,623**
Erhvervsgrad over 2/3				6,311**	4,965**
Har modtaget sygedagpenge				0,891**	0,900**
København				0,864**	1,065*
Vestegnskommuner				1,044+	1,116**
Aarhus				0,887**	0,886**
Aalborg				0,915*	1,000
Odense				0,836**	0,859**
*Indvandrerandel i kom. (kvintil)	0,958**	0,939**	0,991	0,988	1,050**
*Lokal ledighed	0,954**	0,965**	0,946**	0,990+	0,966**
Topleder eller lønmodt. på højeste niveau					1,272**
Øvrige lønmodtagere					0,744**
Selvstændig					0,048**
Ledige (uge 48)					0,587**
Syge- og barselsdagpenge					0,271**
Aktivering og rehabilitering					0,234**
Kontanthjælp og int.ydelse					0,054**
Øvrige uden for arbejdsstyrken					0,141**
Under uddannelse					0,188**
Landbrug					3,289**
Industri					2,490**
Byggeri og anlæg					1,365
Handel og transport					1,586*
Kommunikation og it					3,020**
Finans og forsikring					3,762**
Ejendomshandel og udlejning					2,891**

	Ikke-vestlige indvandrere				
	Reduceret model A			B: A + ledig- hed, er- hvervsgrup- pe, sygedag- penge, kom.	C: Endelig specifikation
	Alle	Mænd	Kvinder		
Erhvervsservice (fx rengøring)					1,525*
Offentlig sektor					2,319**
Kultur og fritidsaktiviteter					1,813**
Militært arbejde					3,415**
Ledelse på øverste niveau					3,882**
Højt specialiseret arbejde					4,664**
Arbejde, der forudsætter viden på mel.niveau					3,684**
Kontorarbejde					2,479**
Detailsalg, service- og omsorgsarbejde					1,469**
Arbejde i de primære erhverv					1,298
Håndværkspræget arbejde					1,861**
Maskin-, transport- og anlægsarbejde					2,855**
Andet arbejde					2,029**
Iran					1,188**
Irak					0,871**
Vietnam					1,245**
Libanon					0,941
Pakistan					0,974
Bosnien					1,580**
Jugoslavien					1,208**
Makedonien					1,561**
Kroatien					1,465**
Serbien-Montenegro					0,982
Thailand					1,300**
Somalia					0,900*
Sri Lanka					1,285**
Filippinerne					1,256**
Marokko					1,268**
Afghanistan					0,991
Kina					0,937
Ukraine					1,690**
Rumænien					1,492**
Indien					1,674**
Rusland					1,322**
Brasilien					1,155*
Bulgarien					1,199*
Syrien					0,808*
Ghana					1,241**
Nepal					0,498**
Øvrige ikke-vestlige lande					1,173**
Højst 4 år siden indrejse	1,495**	1,711**	1,249**	1,430**	0,863**
10-14 år siden indrejse	1,337**	1,223**	1,435**	1,505**	1,098**
15-19 år siden indrejse	1,306**	1,157**	1,450**	1,341**	1,113**
Minimum 20 år siden indrejse	0,876**	0,792**	0,986	0,884**	1,080**
Flygtning eller vurderet flygtning	0,735**	0,818**	0,683**	0,865**	1,002
Lav risiko for sprogproblemer	2,948**	3,119**	2,762**	2,351**	1,014
Mellemhøj risiko for sprogproblemer	0,510**	0,505**	0,491**	0,617**	1,013
Høj risiko for sprogproblemer	0,180**	0,226**	0,166**	0,241**	0,988
Antacid (A02)	0,894**	0,912**	0,871**	0,917**	0,963+
Insulin (A10)	0,822**	0,786**	0,889*	0,805**	0,811**
Antithrombose (B01)	1,486**	1,410**	1,630**	1,329**	0,981
Hjerteterapi (C01)	0,699**	0,703**	0,705**	0,726**	0,827+

	Ikke-vestlige indvandrere				
	Reduceret model A			B: A + ledig- hed, er- hvervsgrup- pe, sygedag- penge, kom.	C: Endelig specifikation
	Alle	Mænd	Kvinder		
Diuretica (C03)	0,959	0,833*	0,959	0,945	1,008
Betablokkere (C07)	0,826**	0,818**	0,831**	0,846**	0,975
Calciumantagonister (C08)	0,955	0,959	0,961	0,950	0,986
Renin-angiotensin syst. (C09)	0,995	1,073	0,907+	0,954	0,978
Lipid-modif. præp. (C10)	0,936+	1,011	0,865*	0,877**	0,964
Antireumatica og antiinflam. (M01)	0,960**	0,953*	0,952*	0,986	0,947**
Analgetica (N02)	0,682**	0,657**	0,699**	0,676**	0,855**
Pyskoleptica og psykoanaleptica (N05-6)	0,460**	0,412**	0,490**	0,473**	0,720**
Anti-afhængighed (N07B)	1,554**	1,347*	2,861**	1,303*	0,668*
KOL-præparater (R03)	1,085**	1,076*	1,076*	1,070**	1,070*
Antihistaminer (R06)	1,127**	1,131**	1,101**	1,112**	1,018
Infektioner (A, B)					1,025
Svulster (C00-D48)					1,022
Blodsygdomme mv. (D50-D89)					0,884
Ernæringssygdomme mv. (E)					1,042
Psykiske lidelser (F)					0,705**
Nervesygdomme (G)					0,950
Øre- og øjensygdomme (H)					1,001
Kredsløbssygdomme (I)					1,016
Åndedrætssygdomme (J)					0,975
Fordøjelsessygdomme (K)					0,974
Hudsygdomme (L)					1,030
Knoglesygdomme mv. (M)					0,847**
Urinvejssygdomme mv. (N)					1,057+
Svangerskabs., misdan. mv. (O, P, Q)					1,064+
Div. symptomer (mavesmerte, besvimelse mv.) (R)					0,922**
Ydre påvirkning (ulykke mv.) (S, T, X, Y)					0,946+
Undersøgelse (Z00-Z13)					0,930**
Kontakt ifm. reproduktion (Z30-Z39)					0,726**
Øvrige forhold (Z20-Z29, Z40-Z99)					0,955
*Alm. læge (antal kontakter)					0,996**
Tandlæge					1,110**
Fysioterapi					1,005
Kiropraktik					1,238**
Psykologhjælp					0,898**
Psykatri					0,673**
Hud- og kønssygdomme					1,006
Øjenlæge					1,008
Ørelæge					0,970
Gynækologi					1,003
Reumatologi					0,916+
Intern medicin					1,089
Kirurgi (plastik, ortopædisk mv.)					1,035
Antal personer	167.801	78.766	89.035	167.801	167.801

Baseline er en gift, 35-39-årig kvinde med to børn. Hun arbejder som ufaglært lønmodtager, har høj erhvervsgrad og har hverken modtaget sygedagpenge eller været ledig. Hvis hun er indvandrer, er hun indvandret fra Tyrkiet og har boet i Danmark i 5-9 år. '+' indikerer, at odds-ratioen er signifikant forskellig fra en på et 10%-niveau, mens '**' og '**' angiver signifikans på hhv. 5% og 1%-niveau.

Bilag 4: Løn, selvforsørgelse og førtidspension blandt ikke-vestlige indvandrere og sammenlignelige danskere

Vi fastslog indledningsvis, at ikke-vestlige indvandrere og danskere adskiller sig med hensyn til deres tilknytning til arbejdsmarkedet, men vi dokumenterede også, at der var en vis forskel i den demografiske, socioøkonomiske og helbredsmæssige sammensætning af de to grupper. For at undersøge, i hvor høj grad forskellen i tilknytningen til arbejdsmarkedet blandt ikke-vestlige indvandrere og danskere kan tilskrives, at danskerne har andre karakteristika end indvandrerne, har vi udført en analyse, som undersøger spørgsmålet for disse to grupper, men ikke for ikke-vestlige efterkommere. Nedenfor præsenterer vi kortfattet resultatet.

Vores grundlæggende problem er, at danskerne er forskellige fra indvandrerne på en række centrale områder, der kan være afgørende for løn mv. Den metode, vi har anvendt, kendes i litteraturen som 'propensity score matching'. Matchingen er udført i Stata med modulet 'PSMATCH2' (se Leuven & Sianesi, 2003). Ideen er, at man kunstigt sammensætter en kontrolgruppe af danskere, der ligner indvandrerne på de karakteristika, som man inkluderer. Vi har lavet en enkelt model, der kontrollerer for alder, køn, civilstand, børn, bopæl, uddannelse, lokal ledighed og indvandrerandel samt køb af medicin og hospitalsdiagnoser. Desuden opdeler vi indvandrerne efter deres estimerede sproglige kompetencer og på deres opholdsgrundlag. Metoden er intuitiv: På baggrund af de observerede variable finder vi en gruppe af danskere, der i gennemsnit ligner indvandrerne med hensyn til de observerede variable. Når vi efterfølgende sammenligner indvandrere og den nye kontrolgruppe, ved vi, at en tilbageværende forskel i løn, selvforsørgelsesgrad og overgangen til førtidspension mellem indvandrere og danskere må skyldes variable, som vi ikke observerer. Det kan eksempelvis være motivation, evner, diskrimination og mange andre forhold.

Kvaliteten af den matchede kontrolgruppe afhænger altså af, i hvor høj grad det er lykkedes at udjævne de observerede forskelle, der faktisk eksisterer, blandt indvandrere og danskere. Når vi fx ved, at der er flere unge blandt indvandrerne, så kan det være årsagen til, at de har lavere løn. Efter matchingen skal denne aldersforskel altså være udjævnet, så gennemsnitsalderen er den samme i begge grupper. Med andre ord skal de observerede variable balancere efter matchingen.

For at balancere observerbare variable må vi først sammensætte den matchede kontrolgruppe af danskere. Vi vil ikke gå i detalje med metoden, men i stedet formidle intuitionen bag. Vi henviser den specielt interesserede læser til den engelsksprogede litteratur (fx Caliendo & Kopeinig 2005). Matching går i sin simpleste variant ud på at finde en dansker, der ligner hver indvandrer i datasættet, dvs. med samme observerbare karakteristika, dvs. køn, alder, uddannelse etc. Hvis listen med relevante observerbare karakteristika er kort og fx kun omfatter køn og alder, vil problemet kunne løses i en simpel krydstabel. Men hvis der er mange variable, så bliver krydstabellen umulig at konstruere, og vi risikerer at få halvtomme celler – dvs. kombinationer af karakteristika, som kun ganske få personer besidder.

I stedet bruger vi de observerede karakteristika til at estimere et indeks. I dette tilfælde er indekset en sandsynlighed, fordi vi estimerer i samme logistiske model, som vi har brugt gennem hele rapporten. Det, vi estimerer, er sandsynligheden for, at en given person er indvandrer. Udfaldsvariablen er altså lig med en, hvis personen er indvandrer, og ellers nul. Vi regresserer på de observerede karakteristika, som vi nævnte ovenfor: Alder, køn, civilstand, børn, bopæl, lokal ledighed og indvandrerandel samt køb af medicin og hospitalsdiagnoser. Uddannelse indgår implicit, fordi vi laver analyserne separat for uddannelseskategorierne 'grundskole', 'faglært' og 'videregående uddannelse'. Vi matcher altså ikke specifikt på den gruppe af indvandrere, der har uoplyst uddannelse, fordi vi ikke kan finde en eksakt match blandt danskerne.

På baggrund af estimerne for de inkluderede variable beregner vi sandsynligheden for, at hver person er en indvandrer (det er her modellens navn – propensity score – kommer fra). For både indvandrere og for danskere har vi nu en sandsynlighed for, at personen er indvandrer, og denne sandsynlighed er betinget på de observerede karakteristika i modellen, dvs. køn, alder, helbred mv. På baggrund af et individs sandsynlighed kan vi 'matche' hver indvandrer med den dansker, der ligner mest. Hver dansker bruges kun én gang, og på den måde skaber vi en ny population af danskere, der en for en er matchet med deres 'indvandrer-tvilling'. Sagt på en anden måde: En indvandrer, der har 25%'s sandsynlighed for at være indvandrer, givet de observerede variable, vil blive matchet med en dansker, der også har 25%'s sandsynlighed for at være indvandrer [*sic!*], målt på de samme variable. Hvis en sådan dansker ikke eksisterer, så vil softwaren vælge den 'nærmeste nabo'. Jo længere væk den danske nabo ligger – målt på sandsynligheden for, at han er indvandrer – des mere forskellig vil han være fra indvandrerens, og desto større vil usikkerheden blive. Hvis usikkerheden bliver for stor, så vil det gå ud over den endelige balancering af de forklarende variable. Det er altså essentielt for analysens validitet, at den nærmeste nabo ligner indvandrerens.

Modellen kan kun inkludere observerbare faktorer, der både er defineret for indvandrere og danskere. Derfor har vi ikke sprog og opholdsgrundlag med for indvandrerne. Vi inkluderer alligevel oplysningerne indirekte ved at opdele indvandrerne efter, om de har få eller mange sprogproblemer (dvs. kun de to ydergrupper), og om de er flygtninge, og køre tre separate modeller – én for hver type af indvandrere. Implicit betyder det også, at vi ikke inkluderer oplysninger, hvor der ikke er overlap mellem danskere og indvandrere, fx erhvervs erfaring. Intuitionen er, at man ikke kan sammenligne individer på faktorer, hvor de er usammenlignelige. Den tekniske forklaring er, at der ikke vil være 'common support' mellem indvandrere og danskere i den underliggende variabel, fx erhvervs erfaring. Vi kan ikke løse problemet ved at opdele samplet, som vi gør med sprog og flygtningestatus, for vi kan under alle omstændigheder ikke skabe sammenlignelighed, når der ikke er basis for det. De separate modeller for sprog og opholdsgrundlag garanterer selvfølgelig ikke, at danskerne ligner på disse karakteristika, da det er umuligt, men de siger derimod noget om graden af heterogenitet i resultaterne.

Tabel B4.1 Balancering af forklarende variable

	Balancering af forklarende variable (absolut bias)						Gns. sandsynlighed for at være indvan- drer (score)		Model fit
	Før matching			Efter matching					
	Gns.	Min.	Maks.	Gns.	Min.	Maks.	Dansk	Indvan- drer	
Grundskole	14,10	0,03	78,13	2,50	0,03	15,00	0,0868	0,2886	0,22
Erhvervsuddannelse	9,41	0,39	67,38	1,43	0,01	4,82	0,0478	0,1217	0,12
Videregående	7,73	0,34	54,29	1,58	0,01	8,30	0,0356	0,0699	0,08
Få sprogproblemer	8,61	0,01	62,23	1,00	0,02	3,38	0,0194	0,0401	0,08
Mange sprogprobl.	14,40	0,23	57,62	1,75	0,03	11,47	0,1558	0,2300	0,37
Flygtninge	10,95	0,62	41,70	1,34	0,03	7,00	0,0261	0,0986	0,14
Alle	10,16	0,12	56,50	1,89	0,05	9,02	0,0600	0,2100	0,19

Før vi ser på resultaterne, skal vi undersøge, om matchingen formår at balancere de observerede karakteristika. Tabel B4.1 viser blandt andet den gennemsnitlige absolutte bias før og efter matchingen (første og fjerde kolonne). Biasen er et udtryk for den gennemsnitlige balancering af alle observerede variable før og efter matchingen. Vi vil ikke komme ind på, hvordan udtrykket er beregnet men blot konstatere, at et gennemsnitligt bias på under 3% anses som værende tilfredsstillende (Caliendo & Kopeinig 2005). Et højt tal angiver, at der er stor forskel på indvandrerne og kontrolgruppen, mens et lavt tal angiver en høj grad af sammenlignelighed. Gennemsnittet er uvægtet, og vi tager således ikke højde for, om nogle variable er vigtigere end andre. Tabellens 7. og 8. kolonne viser den gennemsnitlige prædikterede sandsynlighed for indvandrere og danskere. Hvis den gennemsnitlige sandsynlighed for at være indvandrer er høj for indvandrere, betyder det, at de inkluderede observerede karakteristika er gode til at forklare, om en person er indvandrer eller ej. Modellens fit kan man også aflæse i tabellens sidste kolonne målt ved det såkaldte pseudo R². Hvis fittet er lavt, betyder det ikke, at matchings kvaliteten er dårlig, men snarere at der er vigtige uobserverede forhold, der ikke er med i modellen.

Matchings kvaliteten afhænger derimod kun af, om teknikken formår at balancere de observerede variable, og som vi ser, når vi sammenligner balanceringen før og efter matchingen (kolonne 1 og 4), så sker der en betragtelig reduktion i bias – og for alle modeller gælder, at den gennemsnitlige bias er under 3% efter matching. Tabellens sjette kolonne afslører imidlertid både, at der er variation mellem modellerne, og at enkelte variable er dårligt balancerede. For eksempel ser vi ud for 'Alle' i tabellens sidste række, at selvom den gennemsnitlige bias efter matching kun er på 1,89%, findes der en variabel, hvor bias er 9%. I dette tilfælde drejer det sig om variablen, der indikerer, om personen har fire eller flere hjemmeboende børn. I datasættet er der 6,6% af indvandrerne, der har fire eller flere hjemmeboende børn, mens det gælder 1,3% af den ikke-matched population af danskere. I den matchede kontrolgruppe af danskere er der derimod 4,9%, der har fire eller flere hjemmeboende børn. Selvom vi altså oplever en klart bedre balancering efter matchingen, så er der stadig en statistisk signifikant forskel på netop denne og andre variable. Af pladshensyn har vi valgt ikke at præsentere balanceringsresultaterne for de enkelte variable for alle modeller, men på opfordring kan de rekvireres fra forfatteren. Præmissen for resultaternes validitet er, at vi er villige til at acceptere de kvalitativt små, men statistisk signifikante, afvigelser i balanceringen af enkelte variable. Da den gennemsnitlige balancering er under det kritiske niveau, og fordi afvigelser-

ne er små, anser vi indvandrere og kontrolgruppen af danskere for sammenlignelige på de observerede karakteristika, hvorfor resultaterne er troværdige.

Tabel B4.2 viser resultatet fra matchingen, hvor indvandrere og danskere er opdelt efter deres højeste fuldførte uddannelse. Vi fokuserer på fire udfaldsvariable, som beskriver tilknytningen til arbejdsmarkedet: 1) den gennemsnitlige lønindkomst i første halvår af 2011 (afrundet til tusinde kroner); 2) andelen af personer, som har en lønindkomst på 125.000 kr. eller derover i første halvår af 2011; 3) andelen af selvforsørgede i en tilfældig valgt uge i 2012 (en person anses for at være selvforsørgende, hvis denne ikke modtager en offentlig overførselsindkomst); 4) samt andelen af borgere på førtidspension i samme tilfældige uge i 2012. De to sidstnævnte oplysninger stammer fra DREAM, mens lønnen stammer fra eIndkomstregisteret. Datagrundlaget er det samme, som vi tidligere har benyttet, dvs. personer på arbejdsmarkedet i 2009. De forklarende (observerede) variable relaterer sig også til 2009. Tabel B4.2 viser fordeling af udfaldsvariablene før ('Faktisk') og efter matchingen ('Korrigeret'). Matchingen er udført 1-til-1 på uddannelse for danskere og ikke-vestlige indvandrere, mens Tabel B4.3 viser resultaterne for indvandrere opdelt efter sprog og flygtningestatus samt alle indvandrere samlet. Alle differencer mellem danskere og indvandrere er signifikant forskellige på et 1%'s niveau både før og efter matching.

Tabel B4.2 Propensitiy score matching. Løn, selvforsørgelse og førtidspension blandt indvandrere og sammenlignelige danskere. Opdelt på uddannelse

	Population	Grundskole		Erhvervsuddannelse		Videregående uddannelse	
		Indvan- drere	Danskere	Indvan- drere	Danskere	Indvan- drere	Danskere
Løn (i 1.000 kr.)	Faktisk	60	109	91	146	117	190
	Korrigeret	60	108	91	151	118	195
Løn (>125 t./kr.)	Faktisk	26,7%	51,9%	43,8%	69,6%	50,7%	78,7%
	Korrigeret	26,7%	49,9%	43,8%	69,9%	50,7%	77,9%
Selvforsørgede	Faktisk	49,8%	63,6%	61,9%	79,9%	67,7%	84,3%
	Korrigeret	49,9%	63,8%	61,9%	79,9%	67,7%	83,6%
Førtidspension	Faktisk	5,7%	4,4%	4,6%	1,3%	3,3%	0,7%
	Korrigeret	5,7%	2,9%	4,6%	1,1%	3,3%	0,7%

For personer, der højst har fuldført en grundskoleuddannelse, viser de rå data, at indvandrerne har en gennemsnitlig halvårlig løn på 60.000 kr., mens danskerne i gennemsnit har knap det dobbelte. Men det *faktiske* tal for danskerne kan være farvet, fordi de adskiller sig på mange områder. I tabellens anden række vises den statistisk korrigerede gennemsnitsløn for den matchede kontrolgruppe af danskerne. Her kan vi se, at gennemsnitslønnen kun falder med 1.000 kr., når vi ser på en gruppe af danskere, der er demografisk, uddannelsesmæssigt og helbredsmæssigt sammenlignelige med indvandrerne.¹⁵ Det betyder, at for indvandrere og danskere, der er sammenlignelige på de inkluderede observerede demografiske og helbredsmæssige faktorer, så er der stadigvæk en gennemsnitlig lavere løn blandt indvandrerne. Det betyder dermed, at de observerede karakteristika ikke forklarer, hvorfor indvandrerne har lavere gennemsnitsløn. Vi kan ikke sige, hvilke ikke observerede faktorer der forklarer forskellen, blot at forskellen ikke forsvinder, når vi sammenligner indvandrerne med danskere, der ligner, målt på de inkluderede variable.

Indvandrere med en erhvervsuddannelse tjener i gennemsnit ca. 50% mere end indvandrere med kun en grundskoleuddannelse, men de har stadig en gennemsnitlig løn, der både er lavere end danskere med en grundskoleuddannelse og danskere med en erhvervsuddannelse. Faktisk skal de ikke-vestlige indvandrere have en videregående uddannelse, før deres gennemsnitsløn kan matche den for danskerne med en grundskoleuddannelse. Indvandrere med en videregående uddannelse tjener i gennemsnit 118.000 kr., mens den tilsvarende gruppe af danskere halvårligt tjener 195.000 kr. For alle resultater gælder, at forskellen mellem indvandrere og danskere slet ikke udjævnes, selvom vi korrigerer for de nævnte forhold. Når gennemsnitslønnen for danskerne tværtimod stiger efter matchingen, så skyldes det formentlig, at den matchede kontrolgruppe er lidt yngre end den rå kontrolgruppe, hvilket kan påvirke gennemsnitslønnen.

Denne tendens er den samme, selvom vi i stedet for løn betragter andelen af borgere med en løn på mere end 125.000 kr. Kun 27% af indvandrerne, der højst har fuldført en grundskoleuddannelse, har en løn på 125.000 kr. eller mere, mens det gælder for knap halvdelen af danskerne. I runde tal betyder det, at en indvandrer med en grundskoleuddannelse har en godt 50% lavere chance for at have en løn på 125.000 kr. eller mere i forhold til en dansker, der er sammenlignelig på de observerede karakteristika. For indvandrere med en erhvervsuddannelse eller en videregående uddannelse er odds-ratioen knap 2/3. Konklusionen er, at når vi ser på løn, så kan køn, alder, børn, civilstand, uddannelse, helbred mv. langt fra forklare forskellen på indvandrere og danskere – faktisk, så kan de nævnte variable stort set ikke forklare noget af forskellen.

Kun knap halvdelen af indvandrerne med en grundskoleuddannelse er selvforsørgede, mens det er 64% af de sammenlignelige danskere. For indvandrerne med en erhvervs- eller videregående uddannelse er det henholdsvis 62% og 68%, der er selvforsørgede, mens ande-

¹⁵ Det faktiske tal for indvandrerne kan også ændre sig en smule før og efter matching. Det skyldes, at der findes nogle få indvandrere, for hvem det ikke er muligt at finde en statistisk set sammenlignelig dansker. Alt i alt drejer det sig om mellem 0 og 100 personer, som vi må slette af den årsag. Hvis vi havde benyttet en strengere 'common support'-restriktion end min-max-metoden, så vil antallet af slettede personer stige (større varians), mens balanceringen samtidig ville forbedres (mindre bias).

len for sammenlignelige danskerne er 80% og 84%. Godt 3% af indvandrerne med en videregående uddannelse overgår i perioden 2010 til 2012 (uge 11) til førtidspension, mens det er under 1% af de sammenlignelige danskere. Også her gælder, at de observerede demografiske og helbredsmæssige forskelle langt fra forklarer forskellen.

I Tabel B4.3 giver vi et bud på, hvordan indvandrere i de ekstreme ender af sprogfordelingen klarer sig i forhold til en gruppe af danskere, der er demografisk og helbredsmæssigt sammenlignelige. Vi ser kun på indvandrere i den nederste og øverste fjerdedel af sprogfordelingen, dvs. personer, hvor vi har estimeret, at de kun oplever få eller ingen sprogproblemer samt personer, der selv vurderer, at de har store sprogproblemer. Denne analyse er også udført separat for hver af de to grupper af indvandrere, men modsat resultaterne opdelt på uddannelse, som vi så ovenfor, så er danskerne i denne sammenligningsgruppe *ikke* sammenlignelige på sprog. Formålet er blot at afdække, hvor stor forskel der er i vores udfaldsvariable, når vi ser på indvandrere med få eller mange sprogproblemer. Grunden til, at vi laver denne analyse, er både, at indvandrere med forskellige sprogkompetencer også kan være forskellige i demografisk og helbredsmæssig henseende, hvilket matchingen tager højde for, men samtidig også for at undersøge, om forskellene i udfaldsvariablene mellem indvandrere med dårlige sprogkompetencer og ellers sammenlignelige danskere er større end forskellen mellem sprogmæssigt bedre stillede indvandrere og danskere. Det er imidlertid vigtigt at slå fast, at hvis det er tilfældet, så betyder det *ikke nødvendigvis*, at diskrepansen kun kan tilskrives sprogmæssige barrierer. Det kan også skyldes uobserverede forskelle, der samtidig påvirker både sandsynligheden for at være dårlig til sprog og udfaldet på arbejdsmarkedet.

Resultaterne i tabellen viser, at indvandrere med få estimerede selvvaluerede sprogproblemer har en gennemsnitsløn på 135.000 kr. og knap 61% har en løn over 125.000 kr. De demografisk og helbredsmæssigt sammenlignelige danskere har en gennemsnitsløn på 160.000 kr., mens 67% har en løn over 125.000 kr. Omregnet til odds-ratio, er sandsynligheden for, at indvandrere med få sprogproblemer har en løn på 125.000 kr. eller mere kun 9% lavere end for danskerne. Ser vi i stedet på indvandrere med mange sprogproblemer, så er gennemsnitslønnen 34.000 kr. og kun 12% har en løn på 125.000 kr. eller mere. For de demografisk og helbredsmæssigt sammenlignelige danskere er gennemsnitslønnen på 122.000 kr., mens knap 54% har en løn på 125.000 kr. eller mere. Her er odds-ratioen altså godt 75% lavere for indvandrerne. Det står derfor klart, at indvandrere, der er kendetegnet ved store sprogmæssige udfordringer, også klarer sig relativt dårligt på arbejdsmarkedet, men vi kan ikke nødvendigvis konkludere, at grunden hertil er de sprogmæssige vanskeligheder.

Afslutningsvis skal vi nævne, at tabellen også viser resultater kun for de indvandrere der faktisk er eller vurderes at være flygtet hertil. For denne gruppe, hvor resultatet skal tolkes med de samme forbehold, som vi netop har diskuteret omkring sprog, gælder, at de arbejdsmarkedsmæssigt klarer sig bedre end indvandrerne med dårlige sprogkompetencer, men også at de ligger tættere på gruppen med mange sprogproblemer end på gruppen med få sprogproblemer. Endelig viser tabellen, hvordan indvandrerne set som en helhed klarer sig i forhold til demografisk og helbredsmæssigt sammenlignelige danskere. De ovenstående resultater har allerede fastslået, at der er en betydelig heterogenitet blandt indvandrerne, men samlet set er det godt 34% af indvandrerne, der opretholder en halvårlig løn på 125.000 kr. eller

mere, mens det gælder for 60% af de sammenlignelige danskere. Resultatet før og efter matchingen viser, at hvis vi ikke havde justeret for forskellene i alder, uddannelse, helbred mv., så ville vi have konkluderet, at forskellen er større end den egentlig er.

Sammenfattende kan vi sige, at det ikke er forskelle i demografiske og helbredsmæssige forhold, der er hovedårsagen til forskellen på, hvordan indvandrere med ikke-vestlig herkomst og danskere klarer sig på arbejdsmarkedet. Forskellige demografiske og helbredsmæssige karakteristika er en del af forklaringen, men kun en lille del. Resultaterne indikerer derimod, at sprogmæssige barrierer udgør en større og vigtigere forhindring for en gruppe af indvandrerne, om end vi ikke kan udelukke, at de sprogmæssige indikationer blot er udtryk for en underliggende uobserveret sammenhæng.

Tabel B4.3 Matching; Løn, selvforsørgelse og førtidspension. Opdelt på sprog og opholdsgrundlag

	Population	Få sprogproblemer		Mange sprogproblemer		Flygtninge		Alle	
		Indv.	Dk.	Indv.	Dk.	Indv.	Dk.	Indv.	Dk.
Løn (i 1.000 kr.)	Faktisk	135	159	34	159	74	159	78	159
	Korrigeret	135	160	34	122	74	141	78	140
Løn (>125t./kr.)	Faktisk	60,8%	69,2%	11,9%	69,2%	33,6%	69,2%	34,1%	69,2%
	Korrigeret	60,8%	66,9%	11,9%	53,6%	33,6%	59,7%	34,2%	60,1%
Selvforsørgede	Faktisk	70,8%	77,8%	53,3%	77,8%	47,6%	77,8%	61,2%	77,8%
	Korrigeret	70,8%	76,5%	53,3%	67,2%	47,6%	70,4%	61,2%	71,6%
Førtidspension	Faktisk	1,8%	1,6%	6,7%	1,6%	7,5%	1,6%	4,1%	1,6%
	Korrigeret	1,8%	1,3%	6,7%	2,8%	7,5%	2,4%	4,1%	1,9%

English Summary

Explanatory Analysis of Non-western Immigrants' Labour-market Attachment

In Denmark, non-western immigrants have a much higher probability of receiving income transfers than ethnic Danes. To point out possible mechanisms governing this individually and socially unacceptable relationship, this report investigates the importance of different possible determinants of non-western immigrants' labour-market attachment. We show how different variables correlate with immigrants' probability of at least bringing home an earned income comparable to that of an unskilled worker, but also other outcome measures are considered. All explanatory variables are recovered from administrative registers, so this analysis does not suffer from non-response bias etc. Besides basic demographic variables, we include measures of physical and mental health, educational background, labour-market experience, time of entry and country of origin and an indicator (survey based) of experienced language barriers.

In short, our results consistently show that mental health problems are major risk factors, but also that language barriers are important if the individual has no labour-market experience. And as expected, education is also a both economically and statistically significant correlate of labour-market attachment. The results indicate that refugees from the Middle East or East Africa fare worse than e.g. immigrants from Southeast Asia.



**Det Nationale Institut
for Kommuner og Regioners
Analyse og Forskning**

Købmagergade 22
1150 København K
E-mail: kora@kora.dk
Telefon: 444 555 00