

Aantallen geregistreerde en niet-geregistreeerde burgers uit MOE-landen die in Nederland verblijven

Rapportage schattingen 2009 en 2010

Peter G.M. van der Heijden
Maarten Cruyff
Ger van Gils

Utrecht, 9 januari 2013

Inhoudsopgave:

Managementsamenvatting	i - ii
Hoofdstuk 1: Inleiding	1
Hoofdstuk 2: Methode	3
Hoofdstuk 3: Resultaten	9

Bijlagen:

- 1 - CBS-notitie met aanvullende gegevens voor schatting MOE-landers
- 2 – Sensitiviteitsanalyses
- 3 - Tabellen Geregistreerden en omvangschattingen

Managementsamenvatting

Aantallen geregistreerde en niet-geregistreerde burgers uit MOE-landen die in Nederland verblijven; Rapportage schattingen 2009 en 2010

Peter G.M. van der Heijden (UU), Maarten Cruyff (UU) en Ger van Gils (BOA)

Het voorliggende rapport presenteert schattingen van de omvang van de groepen in Nederland verblijvende personen uit Midden- en Oost-Europa (afgekort als MOE-landers), te weten Polen, Roemenen, Bulgaren en overige MOE-landers¹. Het gaat hier om het aantal geregistreerde MOE-landers en een schatting van het aantal niet geregistreerde MOE-landers. Ook geeft het inzicht in de demografische samenstelling van de groepen en hun geografische spreiding over Nederland.

Medio 2011 is over hetzelfde onderwerp een eerste rapportage uitgebracht over de populatie in Nederland verblijvende MOE-landers in de jaren 2008 en 2009. In dit rapport wordt van dezelfde methoden gebruik gemaakt om tot schattingen te komen en is eenzelfde opzet voor de rapportage gekozen. In het voorgaande rapport ontbraken voor het jaar 2009 gegevens over het totale aantal geregistreerde MOE-landers, omdat gegevens uit het werknemersbestand (WNB) over dat jaar nog niet beschikbaar waren. Nieuw in de onderhavige rapportage zijn daarom de aantallen geregistreerde MOE-landers in 2009 en de aantallen geregistreerde en geschatte MOE-landers in 2010.

De schatting is gedaan op basis van een moderne variant van de capture-recapture methode, waarbij registratiebestanden aan elkaar zijn gekoppeld; te weten de Gemeentelijke Basisadministratie Persoonsgegevens (GBA), de polisadministratie van het UWV (bekend onder de afkorting WNB) en het Herkenningsdienstsysteem van de politie (HKS). Voor het HKS is gekozen omdat dit een van de weinige bestanden is waar niet geregistreerden in voorkomen. Voor de schatting van het aantal niet geregistreerde MOE-landers zijn enkele aannames gemaakt. Twee belangrijke aannamen zijn, ten eerste, dat het percentage HKS-geregistreerden onder de GBA-geregistreerden identiek is aan het percentage HKS-geregistreerden onder de niet-GBA-geregistreerden; ten tweede, dat er geen gemiste koppelingen zijn (van een gemiste koppeling is sprake indien een persoon in GBA bekend is, en in HKS bekend is, zonder dat de koppeling tot stand gebracht wordt).

Op basis van deze aannames is de schatting van het totaal aantal MOE-landers voor 2010 ongeveer 340.000 (met een 95 % betrouwbaarheidsinterval van 322.000 tot 366.000) en voor 2009 305.000 (met een 95 % betrouwbaarheidsinterval van 285.000 tot 325.000) (zie hoofdstuk 3). Deze schattingen hebben echter beperkt belang. De reden hiervan is dat sensitiviteitsanalyses aantonen dat de schattingen erg leunen op de juistheid van de aannamen en het ontbreken van koppelfouten. Geringe afwijkingen van de aannamen in de analyses tot grote veranderingen in de schattingen leiden. Hierbij kan schending van de eerste aanname tot zowel over- als onderschatting van de populatieomvang leiden, en schending van de tweede aanname tot een overschatting.

¹ Het gaat uitsluitend om de migratie uit de genoemde Midden- en Oost Europese landen en niet om migratie uit andere Europese landen.

Het is het meest waarschijnlijk dat, indien de aannamen zijn geschonden, de zojuist genoemde schattingen te hoog zijn, en dus als een maximum dienen te worden opgevat. Voor 2010 is het werkelijke aantal geregistreerden in GBA en WNB ongeveer 182.000. Dit aantal is voor 2010 het minimum. Het werkelijke aantal MOElanders zal in 2010 dus ergens tussen 182.000 en 340.000 liggen.

Vanaf 2008 is er sprake van een gestage toename van alle categorieën geregistreerde MOElanders. De toename is het grootst voor de groep afkomstig uit de overige MOE-landen. De relatieve groei verschilt weinig voor Polen, Roemenen en Bulgaren.

Het beeld voor de niet geregistreerden is anders. De aantallen niet geregistreerde Polen en Bulgaren zijn tussen 2008 en 2010 volgens de schattingen afgenomen (van respectievelijk 55.000 naar 39.000 en van bijna 34.000 naar 25.000). Het aantal niet geregistreerde Roemenen en overige MOE-landers is daarentegen in deze periode naar schatting substantieel (en significant) gegroeid (respectievelijk van ruim 44.000 naar 69.000 en van bijna 37.000 naar 63.000).

Het aandeel van de geregistreerden in de totale populatie MOE-landers lijkt licht te zijn toegenomen (van 49% tot 53%). Polen en Bulgaren zijn beter geregistreerd dan in 2008; Roemenen en overige MOE-landers daarentegen slechter.

De resultaten geven geen aanwijzingen dat het vestigingspatroon van MOE-landers in Nederland zich tussen 2008 en 2010 belangrijk heeft gewijzigd. De belangrijkste woonregio's voor alle groepen blijven de (politieregio's) Haaglanden, Rotterdam-Rijnmond, Amsterdam-Amstellanden en Midden- en West Brabant, waarbij de laatste twee stuivertje wisselen in de top 3.

Vanuit het oogpunt van beleid zijn heeft het voorliggende rapport vooral waarde op de volgende punten: (i) de relatieve geschatte grootte van de vier groepen MOE-landers ten opzichte van elkaar; (ii) informatie over de geregistreerde burgers, waarbij gebruik gemaakt is van de GBA en het werknemersbestand (WNB), en (iii) de samenstelling van de groepen MOElanders in termen van hun achtergrondvariabelen.

Hoofdstuk 1: Inleiding

In de afgelopen tien jaar is de komst van migranten uit de MOE-landen (de Midden- en Oost-Europese landen Polen, Hongarije, Tsjechië, Slowakije, Slovenië, Letland, Litouwen en Estland, Bulgarije en Roemenië) naar Nederland sterk gestegen. Deze stijging is een direct gevolg van de toetreding van Polen, Hongarije, Tsjechië, Slowakije, Slovenië, Letland, Litouwen en Estland tot de Europese Unie in 2004 en van Bulgarije en Roemenië in 2007.

In de afgelopen jaren is in verschillende studies de sociaal-economische positie van met name Polen, Roemenen en Bulgaren goed in beeld gebracht. Voorbeelden van publicaties zijn het Risbo onderzoek *Arbeidsmigranten uit Midden- en Oost-Europa* uit 2009, het *Jaarrapport Integratie 2010* van het CBS en de deelrapportages *Arbeidsmigranten uit Bulgarije, Polen en Roemenië* in 10 gemeentes in Nederland door Godfried Engbersen (EUR) in opdracht van Nicis Institute.

Bekend is hoeveel MOE-landers in de GBA of bij de UWV geregistreerd staan. Waar echter geen zicht op is, is het aantal niet geregistreerde MOE-landers dat in Nederland verblijft.

Het Rijk, maar ook andere partijen die te maken hebben met deze groepen, zoals gemeenten, werkgevers, verhuurders, scholen en welzijnsorganisaties hebben belang bij een goed en betrouwbaar beeld van de daadwerkelijke omvang, samenstelling en locatie van het aantal in Nederland verblijvende geregistreerde en niet geregistreerde groepen MOE-landers.

De onderzoeksvraag is : wat is de omvang, samenstelling en locatie van in Nederland verblijvende geregistreerde en niet-geregistreerde groepen MOE-landers?

Het doel van het onderzoek is om een goede schatting te geven van het daadwerkelijk aantal in Nederland verblijvende MOE-landers en inzicht te krijgen in de demografische samenstelling van de groep en hun geografische spreiding over Nederland.

Het onderzoek onderscheid vier groepen MOE-landers: mensen met een Poolse, Bulgaarse, Roemeense nationaliteit of de nationaliteit van een van de andere Midden- en Oost-Europese landen (Hongarije, Tsjechië, Slowakije, Slovenië, Letland, Litouwen en Estland). Per groep geeft het onderzoek inzicht in de omvang van

- geregistreerde (UWV en GBA) MOE-landers in Nederland; en
- niet-geregistreerde MOE-landers in Nederland, de demografische samenstelling naar:
 - o leeftijd;
 - o geslacht;
 - o geboorteland;
 - o verblijfsduur;
 - o ouderschap; en
 - o burgerlijke staat.

en de geografische spreiding over de 25 politieregio's.

Het onderhavige rapport is de tweede rapportage over de omvang van de populatie MOE-landers in Nederland. Medio 2011 is over hetzelfde onderwerp een eerste rapport uitgebracht. In het huidige rapport wordt van dezelfde methoden gebruik gemaakt om tot schattingen te komen en is eenzelfde opzet voor de rapportage gekozen. In het voorgaande rapport ontbraken voor het jaar 2009 gegevens over het totale aantal geregistreerde MOE-landers.

Nieuw in het huidige rapport zijn daarom de aantallen geregistreerde MOE-landers in 2009 en de aantallen geregistreerde en geschatte MOE-landers in 2010.

Hoofdstuk 2: Methode

2.1 Data

De schatting heeft plaatsgevonden met een moderne variant van de zgn. capture-recapture methode. Bij deze schattingsmethode worden twee of meer bestanden aan elkaar gekoppeld. In het onderhavige geval zijn de volgende bestanden gebruikt:

- GBA gegevens 2009 en 2010
- HKS gegevens 2009 en 2010
- WNB: banenbestand gebaseerd op de Polisadministratie 2009 en 2010.

De GBA en WNB zijn gekoppeld op basis van personen die op de peildatum (laatste vrijdag van september 2009 en 2010) in één van beide bestanden stonden geregistreerd, en de GBA en HKS zijn gekoppeld op basis van personen die in de laatste helft van 2009 en 2010 (01-07 t/m 31-12) in één van beide bestanden stonden geregistreerd. Het koppelen van HKS en WNB is technisch niet mogelijk voor de personen die niet zijn ingeschreven in de GBA.

De schatting is gedaan op basis van nationaliteit, en geboorteland indien nationaliteit onbekend.

Bij koppeling van GBA en HKS, en van GBA en WNB, zijn de volgende basistabellen te maken:

	In GBA	Niet in GBA
In WNB	(1)	(2)
Niet In WNB	(3)	(4)

	In GBA	Niet in GBA
In HKS	(5)	(6)
Niet In HKS	(7)	(8)

Vervolgens kan als volgt een schatting gemaakt worden van de populatie van MOE-landers:

1. koppeling WNB en GBA. In GBA en WNB treffen we louter in Nederland geregistreeerde MOE-landers aan. De som van cellen (1), (2) en (3) is het aantal geregistreeerde MOE-landers, deze som noemen we A.
2. koppeling HKS en GBA. In GBA treffen we louter geregistreeerde MOE-landers aan, maar in HKS treffen we zowel geregistreeerde (in GBA) als niet-geregistreeerde MOE-landers aan. We schatten het aantal MOE-landers bij dat door zowel het HKS als GBA gemist is, cel 8. De som van de cellen (5), (6), (7) en (8) is het totaal aantal geregistreeerde en niet-geregistreeerde personen samen, en we noemen deze som B.
3. Het totaal aantal niet geregistreeerde personen is gelijk aan het verschil $B - A$.

2.2 Het log-lineair model

De schattingen zijn verkregen met behulp van het zgn. log-lineair model. Het log-lineair model is een model waarmee de samenhang tussen categorische variabelen kan worden onderzocht, en bevat zogenaamde hoofdeffecten en interactie-effecten. Het doel van een log-lineaire analyse is een zo spaarzaam mogelijk model te vinden dat zo goed mogelijk op de data past. Van dit type modellen is het onafhankelijkheidsmodel het meest spaarzame model. Het eenvoudigste voorbeeld hiervan is de 2x2 tabel voor GBA en HKS:

	In GBA	Niet in GBA
In HKS	(5)	(6)
Niet In HKS	(7)	(8)

Het model heet een onafhankelijkheidsmodel omdat het veronderstelt dat er geen samenhang is tussen de in het model opgenomen variabelen en enkel uitgaat van hoofdeffecten. In het geval hierboven dus GBA en HKS. Als gevolg daarvan is de verhouding tussen de frequenties in cel 5 en cel 6 gelijk aan de verhouding tussen de frequenties in cel 7 en cel 8.

Het meest uitgebreide log-lineaire model is het verzadigd model; dit model bevat een naast de hoofdeffecten ook interactie-effecten voor elke combinatie van de in het model opgenomen variabelen. Voor het bovenstaande voorbeeld van GBA en HKS betekent de opname van een interactie-effect in het model dat de verhoudingen tussen de frequenties in cel 5 en cel 6 niet meer gelijk hoeven te zijn aan de verhouding tussen de frequenties in cel 7 en cel 8.

Voor deze schatting is de frequentie in cel 8 onbekend en moet dus worden geschat. Om deze cel te kunnen schatten moeten we een aanname maken ten aanzien van de interactie. In de statistiek is het gebruikelijk om in een dergelijk geval van het onafhankelijkheidsmodel uit te gaan, hetgeen dus de afwezigheid van een interactie tussen HKS en GBA impliceert. Zie bijlage 2 voor een overzicht van aannames die ten grondslag liggen aan deze schatting.

Een aanname van het model met alleen de variabelen GBA en HKS is dat alle personen in de populatie een gelijke kans hebben op registratie in GBA enerzijds, en een gelijke kans op

registratie in HKS anderzijds (zie hoofdstuk 2.3.2). Over het algemeen is dit geen realistische aanname, en daarom is het van belang covariaten in het model op te nemen. Met covariaten worden variabelen bedoeld die samenhang vertonen met registratie in GBA en/of HKS, zoals leeftijd en geslacht. Door bijvoorbeeld een interactie-effect tussen HKS en geslacht in het model op te nemen, wordt het mogelijk dat voor vrouwen een andere kans op registratie in HKS wordt geschat dan voor mannen. In de aanwezigheid van covariaten kunnen meerdere interactie-effecten in het model worden opgenomen; het doel van de analyse is dan het meest spaarzame model (een model met zo min mogelijk interactie-effecten) te vinden dat zo goed mogelijk bij de geobserveerde data past.

Hierbij is overigens niet noodzakelijk dat een covariaat in beide bestanden is gemeten. Zwane en van der Heijden (2007) hebben een model ontwikkeld waarbij de ontbrekende waarden op een statistisch verantwoorde wijze worden aangevuld. Hierdoor is het ook mogelijk om schattingen te maken van bijvoorbeeld verblijfsduur, ouderschap en burgerlijke staat, terwijl deze covariaten alleen in GBA zijn gemeten. De aanname die bij het opnemen van een covariaat die alleen in GBA is gemeten dient te worden gemaakt, is dat de ontbrekende waarden niet samenhangen met het wel of niet geregistreerd zijn GBA. Men noemt dit ook wel de MAR ('missing at random') assumptie (zie hoofdstuk 2.3.2).

Er is echter wel een grens aan het aantal covariaten in het model kan worden opgenomen. Het aantal mogelijk interactie-effecten groeit explosief met de toename van het aantal covariaten, en de stabiliteit van de schattingen neemt af naarmate het aantal interactie-effecten in het model toeneemt. Om dit te voorkomen is voor de volgende werkwijze gekozen:

- Als basismodel is het beste model voor de variabelen GBA en HKS, en de covariaten nationaliteit, geslacht, leeftijd en woonregio geselecteerd (deze variabelen zijn zowel in GBA als HKS beschikbaar).
- Het basismodel is gebruikt om de omvang van de MOE-landers populatie te schatten, en het betrouwbaarheidsinterval van deze schatting te bepalen.
- De variabelen verblijfsduur, ouderschap en burgerlijke staat zijn hierna afzonderlijk aan het basismodel toegevoegd om een schatting van de verdeling van deze variabelen in de populatie te krijgen (indien dit een omvangschatting oplevert die afwijkt van de omvangschatting van het basismodel is hiervoor gecorrigeerd).

In dit onderzoek zijn de schattingen dus voor covariaten gecontroleerd. We merken tenslotte op dat, vergeleken met een analyse waarin *niet* voor covariaten wordt gecontroleerd, een analyse waarin *wel* voor covariaten wordt gecontroleerd zowel tot hogere als tot lagere schattingen kan leiden.

2.3 Betrouwbaarheid en bruikbaarheid van de schattingen

De schattingen leveren het meest betrouwbare beeld van de daadwerkelijke omvang, de samenstelling en locatie van het aantal in Nederland verblijvende Polen, Bulgaren, Roemenen en overige MOE-landers dat met de beschikbare data en schattingsmethoden is te geven. De schattingen zijn bruikbaar voor de verschillende partijen die te maken met deze groepen immigranten. Wel willen we een aantal kanttekeningen maken bij de schattingen.

Er zijn bij het maken van de schattingen aannamen gemaakt over de data en over de gebruikte modellen. Deze zullen we hier achtereenvolgens bespreken.

2.3.1 Aannamen over de data

Dubbeltellingen in HKS.

Het CBS heeft ons er op geattendeerd dat er binnen HKS naar alle waarschijnlijkheid nog dubbeltellingen zitten (zie bijlage 1). Personen komen in de HKS als meerdere personen voor, terwijl het in feite om dezelfde persoon gaat. Dit kan gebeuren als uit de gegevens van de persoon bij de politie niet duidelijk is dat het om dezelfde persoon gaat. Er zijn dan meer HKS personen dan in werkelijkheid het geval is. Dit komt onder meer voor als er een proces verbaal is opgemaakt tegen een persoon in verschillende politieregio's.

In dit onderzoek hebben wij, conform het advies van het CBS, het aantal personen in HKS die niet geregistreerd zijn in GBA, teruggebracht met 5 %. Wij maken hierbij de volgende opmerkingen:

- de modelmatige analyses, waarin gebruikt gemaakt wordt van covariaten, zijn uitgevoerd op niet-gecorrigeerde data. Wij weten immers niet welke personen (met welke covariaten) wij dienen te verwijderen uit de data. Van deze modelmatige analyses rapporteren wij in percentages de covariaat-kenmerken van de schattingen in bijlage 3.
- de schattingen van de populatieomvang van de Polen, Roemenen, Bulgaren en andere MOElanders zijn *wel* gecorrigeerd voor deze 5 procent. Deze schattingen zijn te vinden in hoofdstuk 3 van dit rapport.

Koppelbaarheid van personen in HKS en GBA.

Het door ons gebruikte statistische model gaat er van uit dat de gegevens in GBA en in HKS perfect koppelbaar zijn. Van een gemiste koppeling is sprake indien een persoon in GBA aanwezig is, maar dat dit niet uit de HKS gegevens blijkt. Het CBS maakt ons er op opmerkzaam dat deze mogelijkheid aanwezig is (zie bijlage 1). Wij citeren vrij uit bijlage 1: “Van de niet aan de GBA gekoppelde personen uit de HKS is bekend of ze volgens de HKS al dan niet in het buitenland ‘wonen’. Degenen die in het buitenland wonen, zullen terecht niet met de GBA koppelen (en dus in c blijven). Van personen die volgens de HKS-gegevens in Nederland wonen, maar niet koppelen aan de GBA, kan niet altijd met zekerheid gezegd worden of ze terecht niet met de GBA koppelen. Zij kunnen in Nederland verblijven en niet bij een gemeente ingeschreven staan; dan koppelen ze terecht niet met de GBA. Een voorbeeld daarvan zijn werknemers die korter dan 4 maanden in Nederland verblijven. Maar het kunnen ook gemiste koppelingen zijn.

Van de niet-GBA geregistreerden in de HKS in de MOE-landersschatting van 2008 woont volgens de HKS-gegevens 39 procent in Nederland en 61 procent in het buitenland. Daarmee kan het bij maximaal 39 procent van deze groep om onterecht gemiste koppelingen tussen HKS en GBA gaan. Het zou dan bij alle niet-GBA personen uit de HKS met woonland Nederland om gemiste koppelingen moeten gaan. Dit betekent dat maximaal 39 procent van cel (6) naar cel (5) verschuift. In de schatting van 2009 gaat het om 33 procent. Het is voortsnog onbekend welk deel van deze percentages daadwerkelijk gemiste koppelingen betreft. Het percentage potentieel gemiste koppelingen voor 2010 is onbekend. We gaan er van uit dat het van vergelijkbare omvang is als in eerdere jaren.

Bij de MOE-landersschatting kan al wel rekening gehouden worden met de al bekende geregistreerde bevolking, zodat het totale geschatte aantal niet onder het aantal van de geregistreerde bevolking komt.“

Bij hebben hierop het volgende opmerkingen:

- Bij het koppelen van de HKS-gegevens komt het CBS records van – volgens HKS - in Nederland wonende Polen tegen waarin veel informatie ontbreekt. De oorzaak hiervan is bij het CBS onbekend.
- Wij hebben contact opzocht met onze informanten bij de KLPD en de politie. Deze melden ons bij een aanhouding de politie zwaar in zet op identificatie van de aangehoudene. Indien de aangehoudene niet beschikt over papieren, wordt contact gezocht met de ambassade.
- Als dit laatste in de praktijk inderdaad altijd gebeurt, dan is het werkelijke percentage gemiste koppelingen veel lager dan het theoretische maximum van 39 % (voor 2008) en 33 % (voor 2009). Zonder nader onderzoek naar de praktijk bij aanhoudingen valt hierover niet veel te zeggen.
- Het is daarnaast uit ons onderzoek bekend dat er grote aantallen MOE-landers zijn die wel geregistreerd zijn in de WNB als werknemer zonder dat zij ingeschreven staan in de GBA (hiernaar verwijst de laatste alinea in het bovenstaande citaat). Bijvoorbeeld, voor 2008 geldt voor de Polen dat ongeveer 30.000 personen ingeschreven staan in de GBA (waarvan de helft ook in de WNB) en ongeveer 68.000 ingeschreven staan in de WNB *maar niet in de GBA*. Uitgaan van het theoretische maximum van 39 % impliceert dat *alle* personen uit HKS die niet koppelbaar zijn aan de GBA, *niet* ingeschreven staan in de WNB, terwijl juist deze laatste groep veel groter is dan de groep die wel ingeschreven staat in de GBA. Dit vinden wij onaannemelijk en daarmee wordt ook onaannemelijk dat er in de praktijk voor 2008 39 % gemiste koppelingen zijn.

Wij onderzoeken de sensitiviteit van de schattingen voor gemiste koppelingen in bijlage 2.

De aanname van een gesloten populatie.

De aanname van een gesloten populatie impliceert dat er geen in- of uitstroom van MOE-landers plaats heeft gedurende de observatieperiode. Een mogelijke schending is dat onder de personen die wel in HKS voorkomen maar niet in de GBA er veel slechts kort in Nederland verblijven (bijvoorbeeld vakantiegangers, open populatieprobleem). In dat geval is het aantal geschatte personen te hoog.

Om enigszins rekening te houden met het “open” karakter van een migrantenpopulatie zijn de schattingen uitgevoerd op een bestand van een half jaar, en wel de tweede helft van 2008 en de tweede helft van 2009. Het is immers te verwachten dat er vaak leden van deze groepen het land in- en uitreizen. Daarom is het verstandig te werken met gegevens over een zo kort mogelijke periode.

2.3.2 Aannamen van de gebruikte modellen

De onafhankelijkheidsaannname.

De onafhankelijkheidsaannname impliceert dat verhouding tussen personen die wel in HKS voorkomen en die niet in HKS voorkomen hetzelfde is voor personen met een registratie in GBA en personen zonder registratie in GBA. Deze aanname wordt bijvoorbeeld geschonden als de personen die niet in GBA zijn geregistreerd vaker in HKS voorkomen dan personen die wel in GBA zijn geregistreerd. In dat geval leidt de onafhankelijkheidsaannname tot een overschatting van de populatie.

In de rapportage wordt een sensitiviteitsanalyse gerapporteerd, waarbij we een bepaald interactie-effect tussen GBA en HKS veronderstellen en vervolgens kijken wat het effect daarvan op de omvangschatting is.

MAR-assumptie

Bij de opname van variabelen die alleen GBA of HKS zijn gemeten, leidt een schending van de MAR assumptie ('missing at random') tot een minder valide schatting van zowel de populatieomvang en van de samenstelling van de populatie. Om deze reden zijn de variabelen die alleen in GBA zijn gemeten niet gebruikt voor het schatten van de omvang van de populatie, maar alleen voor het schatten van de samenstelling van de populatie op deze variabele.

Hoofdstuk 3: Resultaten

In dit hoofdstuk bespreken we omvangsschattingen. Verdere uitsplitsingen naar achtergrondkenmerken worden besproken in bijlage 3.

In het rapport van medio 2011 ontbraken voor het jaar 2009 gegevens over het totale aantal geregistreerde MOE-landers. Nieuw in het huidige rapport zijn daarom de aantallen geregistreerde MOE-landers in 2009 en de aantallen geregistreerde en geschatte MOE-landers in 2010.

De resultaten voor drie opeenvolgende jaren staan in de tabellen 1 tot en met 3. Onderstaande tabel 1 toont de resultaten voor 2010. De schatting van de Poolse populatie is bijna 170.000 met een 95 % betrouwbaarheidsinterval dat loopt van net onder 160.000 tot ruim 182.000. De telling van het aantal unieke personen in GBA en WNB levert op, dat er ongeveer 130.000 Polen geregistreerd staan in de GBA of de WNB. De schatting van het aantal niet-geregistreerden bedraagt ruim 39.000 Polen (23% van de totale populatie) (namelijk 170.000 – 130.000).

Het geregistreerde deel van de Poolse populatie is percentueel groter dan het geregistreerde deel van de Roemenen, de Bulgaren en de overige MOE-landers. Voor de Roemenen is het geschatte aantal bijna 70.000, waarvan 13 % is geregistreerd; voor de Bulgaren is het geschatte aantal bijna 39.000, maar een groter deel, te weten 36% is geregistreerd; voor de overige MOE-landers is het geschatte aantal ruim 63.000, waarvan 45 % is geregistreerd.

Het geschatte totale aantal MOE-landers bedraagt ongeveer 340.000, waarvan ruim de helft is geregistreerd in de GBA of de WNB.

Tabel 1: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2010

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatieschatting	%	95% betrouwbaarheidsinterval
Pools	130.277	77%	39.482	23%	169.759	50%	(159.976 – 182.707)
Roemeens	9.223	13%	60.002	87%	69.225	20%	(62.693 – 77.692)
Bulgaars	13.773	36%	25.011	64%	38.784	11%	(34.374 – 44.234)
Overig MOE	28.687	45%	34.640	55%	63.327	19%	(58.856 – 69.116)
Totaal	181.960	53%	159.134	47%	341.094	100%	(322.244 – 366.057)

Tabel 2: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2009

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatieschatting	%	95% betrouwbaarheidsinterval
Pools	115.538	81%	26.702	19%	142.240	47%	(133.725 - 152.757)
Roemeens	8.111	13%	55.601	87%	63.712	21%	(57.951 - 70.676)
Bulgaars	12.452	31%	27.250	69%	39.702	13%	(35.113 - 45.926)
Overig MOE	23.007	40%	34.496	60%	57.503	100%	(52.925 - 63.132)
Totaal	159.108	52%	144.050	48%	303.158	100%	(286.367 - 325.735)

Tabel 3: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2008

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatie-schatting	%	95% betrouwbaarheidsinterval
Pools	101.011	65%	55.079	35%	156.090	55%	(144.482 - 169.822)
Roemeens	7.225	16%	37.365	84%	44.590	16%	(39.616 - 50.589)
Bulgaars	10.448	24%	33.670	76%	44.118	16%	(38.656 - 51.060)
Overig MOE	18.562	51%	18.079	49%	36.641	13%	(33.201 - 40.472)
Totaal	137.246	49%	144.194	51%	281.440	100%	(262.390 - 305.031)

In totaal is de geschatte populatie MOE-landers in 2010 ten opzichte van 2009 toegenomen met ca. 38.000 personen, waarvan er ongeveer 23.000 zijn geregistreerd. De Poolse populatie is verantwoordelijk voor een belangrijk deel van deze toename. Het aantal Polen nam toe met ruim 27.000. Meer dan de helft daarvan, ongeveer 15.000, is geregistreerd. Het aantal Roemenen in het land nam toe met ruim 5.000, waarvan ruim 1.000 personen zijn geregistreerd. Het aantal Bulgaren daalde per saldo met ongeveer 1.000 personen, maar meer Bulgaren lieten zich registreren; hun aantal nam met ruim 1.000 toe. Het aantal personen afkomstig uit overige MOE-landen, tot slot, nam met bijna 6.000 toe, terwijl het aantal geregistreerden van deze groep met een nagenoeg zelfde aantal steeg.

De cijfers laten zien dat er vanaf 2008 sprake is van een gestage toename van alle categorieën geregistreerde MOE-landers. Het totale aantal groeit met ongeveer 44.000 van ruim 137.000 tot bijna 182.000. De toename is het grootst voor de groep afkomstig uit de overige MOE-landen. De relatieve groeicijfers lopen voor de Polen, Roemenen en Bulgaren weinig uit een. Het beeld voor de niet geregistreerden wijkt hier van af. De aantallen niet geregistreerde Polen en Bulgaren zijn tussen 2008 en 2010 volgens de schattingen afgenomen (van respectievelijk 55.000 naar 39.000 en van bijna 34.000 naar 25.000). Het aantal niet geregistreerde Roemenen en overige MOE-landers is daarentegen in deze periode naar schatting substantieel (en significant) gegroeid (respectievelijk van ruim 44.000 naar 69.000 en van bijna 37.000 naar 63.000). Het resultaat is een betrekkelijk bescheiden toename van het totaal aantal niet geregistreerde MOE-landers van ruim 144.000 in 2008 tot ruim 159.000 in 2010.

Het aandeel van de geregistreerden in de totale populatie MOE-landers lijkt licht te zijn toegenomen (van 49% tot 53%). Vooral de hogere registratiegraad van de Polen is hiervoor verantwoordelijk. Ook de Bulgaren hebben zich beter laten registreren. De Roemenen en overige MOE-landers daarentegen zijn in mindere mate geregistreerd dan in 2008.

De resultaten geven geen aanwijzingen dat het vestigingspatroon van MOE-landers in Nederland zich tussen 2008 en 2010 belangrijk heeft gewijzigd. De belangrijkste woonregio's voor alle groepen blijven de (politieregio's) Haaglanden, Rotterdam-Rijnmond, Amsterdam-Amstellanden en Midden- en West Brabant, waarbij de laatste twee stuivertje wisselen in de jaarlijkse top 3. Amsterdam-Amsteland is vooral een belangrijke regio voor Roemenen en overige MOE-landers en in iets mindere mate voor Bulgaren, terwijl Midden – en West Brabant belangrijker is voor Polen. Deze patronen gelden zowel voor geregistreerden als niet geregistreerden en veranderen niet of nauwelijks tussen 2008 en 2009.

Bijlagen

Toelichting op bijlage 1: bijlage 1 is een commentaar en aanvulling van de zijde van CBS bij het rapport van 2011. De opmerkingen gelden ook voor de onderhavige nieuwe rapportage. Volledigheidshalve nemen we de bijlage ook op in deze rapportage.

Bijlage 1: CBS-notitie met aanvullende gegevens voor schatting MOE-landers

Aan: BZK
Van: Rik van der Vliet, Bart Bakker en Fred Gast
Datum: 7 juli 2011

De groep van prof. P. van der Heijden van de Universiteit Utrecht (UU) heeft een schatting gemaakt van het totaal aantal MOE-landers in Nederland. Het gaat daarbij niet alleen om de MOE-landers van wie bekend is dat ze in Nederland verblijven doordat ze bijvoorbeeld ingeschreven staan bij de gemeente of geregistreerd staan als werknemer. Ook het aantal MOE-landers dat nergens officieel geregistreerd staat, maar wel in Nederland verblijft wordt geschat.

Anders dan de geregistreerde MOE-landers, kunnen de niet-geregistreerde MOE-landers niet geteld worden. Juist omdat ze niet geregistreerd staan is hun aantal onbekend. Er is echter een registratie waarin deze personen toch voorkomen, namelijk de verdachtenregistratie (HKS). Immers, personen die in Nederland verblijven, kunnen een misdrijf begaan en aangehouden worden, ongeacht of ze ergens officieel geregistreerd staan.

De UU heeft een statistische methode toegepast (gebaseerd op capture-recapture principes) waarmee het totaal aantal niet-geregistreerde MOE-landers geschat kan worden op basis van degenen die in de HKS geregistreerd staan. Daarvoor wordt de HKS-registratie gekoppeld met de Gemeentelijke Basisadministratie (GBA), waarin alle officieel geregistreerde inwoners van Nederland staan. Door het koppelen van de twee bestanden is van alle personen in de GBA bekend of ze ook in de HKS voorkomen. Door nu aan te nemen dat niet-geregistreerden relatief net zo vaak in de HKS voorkomen, en doordat bekend is hoeveel niet-geregistreerden er in de HKS voorkomen, kan het totaal aantal niet-geregistreerden geschat worden. Omdat niet elke persoon gelijke kansen zal hebben om in de HKS voor te komen, corrigeert de methode nog voor verschillen in een aantal achtergrondkenmerken, zoals leeftijd en geslacht. Daarbij blijft wel de aanname gelden dat een niet-geregistreerde persoon net zoveel kans heeft om in de HKS voor te komen als een geregistreerd persoon met dezelfde (bekende) achtergrondkenmerken.

Naast de genoemde algemene aanname van de gehanteerde methode is er impliciet nog een aanname. Het betreft de aanname dat de GBA gegevens en de HKS gegevens volledig met elkaar koppelen. Er mogen geen gemiste koppelingen zijn. Er is echter wel degelijk kans op gemiste koppelingen. In het geval van de koppeling van GBA en HKS zijn er in elk geval twee mogelijkheden te onderscheiden waardoor de koppeling niet volledig is:

- A. Personen uit de HKS kunnen ten onrechte niet teruggevonden worden in de GBA. Dit speelt als de gegevens in de HKS niet juist zijn, waardoor de koppeling met de GBA misloopt.
- B. Personen komen in de HKS als meerdere personen voor, terwijl het in feite om dezelfde persoon gaat. Dit kan gebeuren als uit de gegevens van de persoon bij de politie niet duidelijk is dat het om dezelfde persoon gaat. Er zijn dan meer HKS personen dan in werkelijkheid het geval is. Dit komt onder meer voor als er een proces verbaal is opgemaakt tegen een persoon in verschillende Politie regio's.

Voor beide problemen geldt dat onbekend is in welke mate dit precies speelt, maar dat wel aangegeven kan worden in welke mate het zou kunnen spelen en wel als volgt:

Ad A) Van de niet aan de GBA gekoppelde personen uit de HKS (vakje c, zie Bijlage) is bekend of ze volgens de HKS al dan niet in het buitenland 'wonen'. Degenen die in het buitenland wonen, zullen terecht niet met de GBA koppelen (en dus in c blijven). Van personen die volgens de HKS-gegevens in Nederland wonen, maar niet koppelen aan de GBA, kan niet altijd met zekerheid gezegd worden of ze terecht niet met de GBA koppelen. Zij kunnen in Nederland verblijven en niet bij een gemeente ingeschreven staan; dan koppelen ze terecht niet met de GBA. Een voorbeeld daarvan zijn werknemers die korter dan 4 maanden in Nederland verblijven. Maar het kunnen ook gemiste koppelingen zijn.

Van de niet-GBA geregistreerden in de HKS in de MOE-landersschatting van 2008 woont volgens de HKS-gegevens 39 procent in Nederland en 61 procent in het buitenland. Daarmee kan het bij

maximaal 39 procent van deze groep om onterecht gemiste koppelingen tussen HKS en GBA gaan. Het zou dan bij alle niet-GBA personen uit de HKS met woonland Nederland om gemiste koppelingen moeten gaan. Dit betekent dat maximaal 39 procent van c naar a verschuift (zie Bijlage). In de schatting van 2009 gaat het om 33 procent. Het is vooralsnog onbekend welk deel van deze percentages daadwerkelijk gemiste koppelingen betreft.

Bij de MOE-landersschatting kan al wel rekening gehouden worden met de al bekende geregistreerde bevolking, zodat het totale geschatte aantal niet onder het aantal van de geregistreerde bevolking komt.

Ad B) Vanuit de eerste leveringen van de HKS-data is bekend dat personen die aan de GBA gekoppeld kunnen worden vaker voorkomen als verdachte. Hiervoor wordt in het vervolg van het proces gecorrigeerd. In 2008 was dit voor 5 procent van de wel in GBA geregistreerden het geval. Deze dubbeling was niet meer aanwezig in de bestanden die gebruikt zijn voor de MOE-landers schatting. Personen die niet aan de GBA koppelen kunnen niet geïdentificeerd worden en daarmee kan ook moeilijk gecorrigeerd worden voor dubbel voorkomen. Ervanuit gaande dat er onder HKS-personen die niet in de GBA geregistreerd staan net zoveel dubbelingen voorkomen als in eerste instantie onder de wel GBA-geregistreerden, schatten wij c (zie bijlage) 5 procent lager. In de MOE-landers schatting van 2009 gaat het nog maar om 0,3 procent dubbelingen onder de GBA-geregistreerden. Blijkbaar zijn tussen 2008 en 2009 de wel GBA-geregistreerden in de bronbestanden beter geïdentificeerd waardoor dubbelingen zijn voorkomen. Het is echter niet waarschijnlijk dat dit ook mogelijk is geweest voor niet GBA-geregistreerden die moeilijker identificeerbaar zijn. Daarom gaan we ook in 2009 uit van 5 procent dubbelingen onder de niet GBA-geregistreerden in de HKS.

Situaties A en B hebben invloed op de totale populatieschatting. Als A speelt, zou het aantal HKS-personen dat niet in de GBA geregistreerd staat kleiner moeten zijn en het aantal HKS-personen dat wel in de GBA geregistreerd staat groter. Bij B zou het aantal HKS-personen dat niet in de GBA geregistreerd staat kleiner moeten zijn. In beide gevallen resulteert dat in minder bijgeschatte niet-geregistreerde MOE-landers.

De precieze gevolgen voor de geschatte aantallen moeten met het schattingsmodel bepaald worden. Daarmee zullen de gevolgen van A groter zijn naarmate van een hoger percentage gemiste koppelingen uit wordt gegaan. De grenzen tussen het minimale en maximale aantal mogelijk gemiste koppelingen liggen nu nog ver uit elkaar. Het verdient aanbeveling dit nauwkeuriger af te bakenen. Op dit moment is dit nog niet mogelijk gebleken.

Schattingsmodel:

	HKS	
GBA	<i>ja</i>	<i>nee</i>
<i>ja</i>	a	b
<i>nee</i>	c	d

Bijgeschatte niet-GBA, niet HKS-populatie (d):

$$d = \frac{b * c}{a}$$

Bijlage 2: Sensitiviteitsanalyses

Sensitiviteitsanalyses hebben als doel na te gaan hoe groot de invloed is van aannames die gemaakt zijn om tot de schattingen te komen. In hoofdstuk 2.2 zijn vijf aannames besproken. We vatten ze hier kort samen:

- geen dubbeltellingen in de personen in HKS die *niet* bekend zijn in GBA. Dit probleem hebben wij aangepakt door bij de schattingen in hoofdstuk 3 uit te gaan van 5 % dubbeltellingen, hierbij een advies van het CBS volgend. Overigens is in de andere cellen het probleem van dubbeltellingen nihil, door de hoge kwaliteit van de GBA.
- gemiste koppelingen. Wij onderzoeken de gevoeligheid van de uitkomsten voor gemiste koppelingen in paragraaf 2.
- de aanname van een gesloten populatie. Wij hebben dit probleem gepoogd zo klein mogelijk te maken door de populaties van groepen MOE-landers slechts voor de periode van een half jaar te bestuderen.
- de onafhankelijkheidsaanname. In paragraaf 1 onderzoeken wij de gevoeligheid van de uitkomsten voor afhankelijkheid tussen het geregistreerd zijn in de GBA en het geregistreerd zijn in HKS.
- de MAR-assumptie ('missing at random'). Deze is door ons omzeild door geen modellen te schatten die gebruik moeten maken van deze assumptie.

In paragraaf 3 komen we kort terug op de sensitiviteit van de uitkomsten voor de gemaakte aannames.

1 Onafhankelijkheid tussen registratie in GBA en registratie in HKS.

De onafhankelijkheidsaanname impliceert dat verhouding tussen personen die *wel* in HKS voorkomen en die *niet* in HKS voorkomen hetzelfde is voor personen *met* een registratie in GBA en personen *zonder* een registratie in GBA. Het gehanteerde model gaat uit van onafhankelijkheid, de standaardaanpak bij dit type analyses.

Deze aanname wordt geschonden als de personen die niet in GBA zijn geregistreerd vaker verdachte zijn van een delict (en dus vaker in HKS voor te komen) dan personen die wel in GBA zijn geregistreerd, of andersom, als personen die wel in GBA zijn geregistreerd vaker in HKS voorkomen dan personen die niet in GBA voorkomen. Hoewel wij dit niet eerder in statistische literatuur zijn tegengekomen, is een sensitiviteitsanalyse op dit punt niet ingewikkeld uit te voeren. In deze bijlage bespreken we de uitkomsten van zo'n sensitiviteitsanalyse.

De aanpak is als volgt. In het loglineaire model geldt dat, conditioneel op de gebruikte covariaten nationaliteit, geslacht, leeftijd en woonregio, de kans op registratie in GBA en de kans op registratie in HKS onafhankelijk zijn. In het loglineaire model impliceert dit dat de parameter voor de relatie tussen GBA en HKS nul is. Wij gaan nu schattingen produceren waarin wij deze parameter een waarde geven die afwijkt van nul.

Deze parameter is te interpreteren als een zgn. *conditionele odds ratio* (waarbij dit woord 'conditioneel' verwijst naar de covariaten waarmee in de schatting wordt rekening gehouden. In een tabel met cellen (5), (6), (7) en (8) (vgl. hoofdstuk 2.2) is de odds ratio bij onafhankelijkheid gelijk aan $1 = (5)/(7) / (6)/(8)$, met andere woorden, in GBA is de

verhouding tussen wel of niet voorkomen in HKS identiek aan dezelfde verhouding in niet-GBA. Bij kleine zgn. inclusiekansen² is de odds ratio bijna gelijk aan het zgn. *relatieve risico*, dat mogelijk gemakkelijker te interpreteren is. Bij een relatief risico van 0,8 geldt dat voor personen die wel in GBA zijn ingeschreven de kans op voorkomen in HKS 0,8 keer zo groot is als dezelfde kans voor personen die niet in GBA zijn ingeschreven. Omgekeerd geldt dan dat voor personen die niet in GBA zijn ingeschreven de kans om in HKS voor te komen 1,25 maal zo groot is als voor personen die wel in GBA zijn ingeschreven.

Wij kiezen in de sensitiviteitsanalyse voor de waarden 0,8 (implicierend dat de kans dat personen die *niet* GBA voorkomen, maar wel in HKS 0,8 keer zo groot is als voor personen die niet in GBA voorkomen) en 1,25 (omgekeerde verband). Wij hebben hierin geen onderscheid tussen nationaliteiten aangebracht. Deze waarden zijn tot op zekere hoogte willekeurig gekozen en reflecteren een voldoende substantieel verband, maar, zoals we direct in de tabellen van 2008 kunnen zien, is het kiezen van waarden die veel lager zijn dan 0,8 niet mogelijk omdat we dan in conflict komen met het aantal officieel geregistreerde Polen.

Tabellen 3a, 3b en 3c hebben betrekking op 2010. De gegevens uit tabel 3b komen overeen met die uit tabel 1, en deze zijn hier opnieuw gegeven om de vergelijkbaarheid met de voorgestelde schattingen te vergemakkelijken.

De gegevens over 2009 staan in de tabellen 3d tot en met 3f.

² Hier is de inclusiekans in HKS klein. Voor de relatie tussen de odds ratio en het relatieve risico, zie Agresti (2002), *Categorical data analysis*. Wiley.

Tabel 3a: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2010; OR=0,8

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatie-schatting	%
Pools	130.277	89%	15.395	11%	145.672	100%
Roemeens	9.223	16%	47.949	84%	57.172	100%
Bulgaars	13.773	41%	20.008	59%	33.781	100%
Overig MOE	28.687	53%	25.845	47%	54.532	100%
Totaal	181.960	62%	109.196	38%	291.156	100%

Tabel 3b: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2010; OR=1,0

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatie-schatting	%
Pools	130.277	77%	39.482	23%	169.759	100%
Roemeens	9.223	13%	60.002	87%	69.225	100%
Bulgaars	13.773	36%	25.011	64%	38.784	100%
Overig MOE	28.687	45%	34.640	55%	63.327	100%
Totaal	181.960	53%	159.134	47%	341.094	100%

Tabel 3c: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2010; OR=1,25

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatieschatting	%
Pools	130.277	65%	69.499	35%	199.776	100%
Roemeens	9.223	11%	75.069	89%	84.292	100%
Bulgaars	13.773	31%	31.254	69%	45.027	100%
Overig MOE	28.687	39%	45.633	61%	74.320	100%
Totaal	181.960	45%	221.454	55%	403.414	100%

Tabel 3d: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2009; OR=0,8

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatie-schatting	%
Pools	115.538	95%	6.501	5%	122.039	100%
Roemeens	8.111	15%	44.419	85%	52.530	100%
Bulgaars	12.452	36%	21.798	64%	34.250	100%
Overig MOE	23.007	47%	26.149	53%	49.156	100%
Totaal	159.108	62%	98.867	38%	257.975	100%

Tabel 3e: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2009; OR=1,0

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatie-schatting	%
Pools	115.538	81%	26.702	19%	142.240	100%
Roemeens	8.111	13%	55.601	87%	63.712	100%
Bulgaars	12.452	31%	27.250	69%	39.702	100%
Overig MOE	23.007	40%	27.250	60%	57.503	100%
Totaal	159.108	49%	144.050	48%	303.158	100%

Tabel 3f: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2009; OR=1,25

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatieschatting	%
Pools	115.538	54%	51.953	31%	167.491	100%
Roemeens	8.111	13%	69.579	90%	77.690	100%
Bulgaars	12.452	20%	34.066	73%	46.518	100%
Overig MOE	23.007	44%	44.930	66%	67.937	100%
Totaal	159.108	41%	200.528	56%	359.636	100%

In de tabellen 3a tot en met c zien we dat een wijziging van de odds ratio van 1,0 naar 0,8 of 1,25 een grote impact op de schattingen heeft. Is bij een odds ratio van 1,0 de schatting van alle MOE-landers samen ongeveer 340.000, bij een odds ratio van 0,8 daalt die schatting naar ongeveer 290.000 en bij een odds ratio van 1,25 stijgt de schatting naar ruim 400.000.

Wij merken op dat de odds ratio niet veel lager gezet kan worden dan 0,8: voor de odds ratio van 0,8 is de schatting van het aantal Polen in 2010 ruim 145.000 (ongeveer 24.000 lager dan de schatting bij een odds ratio van 1,0) waarvan ongeveer 130.000 Polen in GBA of de WNB zijn geregistreerd. Bij een daling naar bijvoorbeeld 0,6 wordt de schatting weer ongeveer 24.000 lager en in dit geval komt de schatting van het totaal aantal Polen met ongeveer 120.000 onder het officieel aantal geregistreerden van ruim 130.000 uit.

Wij maken nog de volgende opmerking. De reden dat de schattingen sensitief zijn voor afwijkingen in de odds ratio is gelegen in de relatief geringe overlap tussen GBA en HKS (het aantal in HKS geregistreerde MOE-landers is laag). Nemen we de tabel die weergegeven is in hoofdstuk 2.2, dan is de ontbrekende waarde (8) te schatten als $(8) = (6) \cdot (7) / (5)$. Indien (5) ten opzichte van (6) en (7) relatief klein is, dan wordt (8) snel erg groot en instabiel in de zin dat kleine veranderingen in (5) tot grote veranderingen in (8) leiden. Hiermee wordt dan de som $(5) + (6) + (7) + (8)$ erg instabiel. Voor de MOE-landers speelt dit probleem in hoge mate (zie bijlage 3, tabel A1).

2 Mogelijk gemiste koppelingen

In hoofdstuk 2.4.1 hebben wij het probleem van mogelijk gemiste koppelingen besproken. Voor de personen die in HKS zitten maar niet in GBA, geldt dat zij niet koppelbaar waren. In de CBS notitie (zie bijlage 1) wordt aangegeven dat dit probleem niet speelt voor personen waarvan kon worden vastgesteld dat zij in het buitenland woonachtig waren. In principe zou, aldus de CBS notitie, iedere persoon waarvan in HKS staat dat hij in Nederland woonachtig is, een gemiste koppeling kunnen zijn. Zoals uiteengezet in hoofdstuk 2.4.1 acht de politie dit niet waarschijnlijk omdat zij zegt zwaar in te zetten op identificatie van verdachten, maar evengoed zijn de records van deze personen vaak slecht gevuld.

Het CBS geeft aan dat in 2008 39 % van de personen die in HKS zitten maar niet in GBA, aangeven in Nederland woonachtig te zijn. In 2009 is dit 33 %. Wij hebben de sensitiviteit van de schattingen voor gemiste koppelingen op twee manieren onderzocht: door (vrij willekeurig) aan te nemen dat het percentage 10 % is (dus dit is ongeveer een kwart van 39 %) en door aan te nemen 39 % evenredig verdeeld kan worden over de combinaties HKS-wel GBA en HKS-niet GBA.

Het percentage potentieel gemiste koppelingen voor 2010 is onbekend. We gaan er van uit dat het van vergelijkbare omvang is als in eerdere jaren.

We bespreken eerst de resultaten voor 10 % gemiste koppelingen. Voor 2008, zie tabel 6a voor de resultaten en tabel 6b voor 2009. We zien dat in 2008 bij 10% gemiste koppelingen de totale schatting daalt van ongeveer 280.000 naar ongeveer 195.000; voor 2009 zien we vergelijkbare resultaten. Indien we kijken naar de Polen in Tabel 6a, dan zien we dat 10% waarschijnlijk te hoog is omdat het percentage niet-geregistreerden bij 10% miskoppelingen slechts 6 %. Dat lijkt irrealistisch.

Evengoed moeten we, net als in paragraaf 1 van deze bijlage, concluderen dat de schattingen zeer gevoelig zijn voor miskoppelingen, en dat de schattingen besproken in hoofdstuk 3 een maximum aangeven van het aantal MOE-landers.

Tabel 6a: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2008 bij 10 % gemiste koppelingen

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatie-schatting	%
Pools	101.011	94%	6.909	6%	107.920	100%
Roemeens	7.225	29%	17.545	71%	24.770	100%
Bulgaars	10.448	30%	23.847	70%	34.295	100%
Overig MOE	18.562	63%	10.894	37%	29.456	100%
Totaal	137.246	70%	59.195	30%	196.441	100%

Tabel 6b: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2009 bij 10% gemiste koppelingen

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatie-schatting	%
Pools	115.538	108%	-8.175	-8%	107.363	100%
Roemeens	8.111	25%	24.634	75%	32.745	100%
Bulgaars	12.452	37%	20.923	63%	33.375	100%
Overig MOE	23.007	54%	19.575	46%	42.582	100%
Totaal	159.108	74%	56.958	26%	216.066	100%

We hebben de sensitiviteit ook onderzocht onder een andere aanname, namelijk dat de 39% (in 2008, en 33 % in 2009) in Nederland woonachtige personen aangetroffen in HKS proportioneel verdeeld worden over de cellen HKS-wel GBA en HKS-niet GBA, waarbij dan in deze laatste cel alleen de in het buitenland woonachtigen zijn opgenomen. Zie voor 2008 tabel 7 en voor 2009 tabel 8. De resultaten zijn vergelijkbaar met die van de eerste methode.

Tabel 7: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2008 bij proportionele verdeling van 39% gemiste koppelingen

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatie-schatting	%
Pools	101.011	94%	7.213	6%	108.224	100%
Roemeens	7.225	29%	22.455	71%	29.680	100%
Bulgaars	10.448	30%	20.399	70%	30.847	100%
Overig MOE	18.562	63%	8.694	37%	27.256	100%
Totaal	137.246	70%	58.761	30%	196.007	100%

Tabel 8: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2009 bij proportionele verdeling van 33% gemiste koppelingen

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatie-schatting	%
Pools	115.538	106%	-6.987	-6%	108.551	100%
Roemeens	8.111	18%	36.976	82%	45.087	100%
Bulgaars	12.452	41%	18.170	59%	30.622	100%
Overig MOE	23.007	53%	20.579	47%	43.586	100%
Totaal	159.108	70%	68.739	30%	227.847	100%

3 Discussie en conclusie

De sensitiviteitsanalyses laten zien dat de schattingen gepresenteerd in hoofdstuk 3 erg leunen op de juistheid van de aannamen van onafhankelijkheid van de twee registraties (conditioneel op de covariaten) en het ontbreken van koppelfouten. Relatief geringe afwijkingen van de aannamen leiden tot relatief grote aanpassingen in de schattingen.

Allereerst de eerste aanname: onafhankelijkheid van de twee registraties. Afhankelijkheid van de twee registraties kan leiden tot een lager aantal MOE-landers of een hoger aantal MOE-landers. Dat wil zeggen, als de werkelijke conditionele odds ratio lager is dan 1, zodat personen in GBA relatief gezien minder vaak voorkomen in HKS dan personen niet in GBA, dan is de huidige schatting in hoofdstuk 3 te hoog. We merken op dat voor het jaar 2009 de ondergrens wordt beperkt door kennis over het aantal officieel geregistreerden. Indien de werkelijke conditionele odds ratio hoger is dan 1 dan zijn de schattingen in hoofdstuk 3 te laag.

Uit paragraaf 2 van deze bijlage blijkt dat, indien er sprake is van gemiste koppelingen, de schattingen in hoofdstuk 3 te hoog zijn.

We concluderen dat, indien men bereid is aan te nemen dat de werkelijke conditionele odds ratios kleiner of gelijk zijn aan 1, dat dan de werkelijke populatieomvangen liggen tussen de schattingen in hoofdstuk 3 en het aantal geregistreerden. Dit is nogal een grote range, waardoor de informatie in hoofdstuk 3 waarschijnlijk een beperkte beleidsmatige waarde heeft. We merken verder nog op dat de effecten van afwijkingen van de twee aannamen niet optelbaar zijn (er kan niet sprake zijn van zowel een werkelijke conditionele odds ratio van 0,8 als een percentage gemiste koppelingen) omdat de aangepaste schattingen dan lager worden dan het bekende aantal geregistreerden.

Vanuit het oogpunt van beleid zijn heeft het voorliggende rapport vooral waarde op de volgende punten:

- de relatieve geschatte grootte van de vier groepen MOE-landers ten opzichte van elkaar;
- informatie over de geregistreerde burgers op basis van de gekoppelde gegevens uit GBA en WNB;
- de relaties met covariaten, gegeven in bijlage 3.

Bijlage 3: Tabellen Geregistreerden en omvangschattingen

De volgende tabellen zijn in dit rapport opgenomen:

A. Overzichtstabellen totaal aantal geregistreerde MOE-landers in 2009 en 2010

Tabel A1: Aantal in GBA, HKS en WNB geregistreerde MOE-landers in 2009 naar nationaliteit

Tabel A2: Aantal in GBA, HKS en WNB geregistreerde MOE-landers in 2010

B. Overzichtstabellen van MOE-landers met een registratie in GBA in 2010

Tabel B1: Registraties nationaliteit en geslacht in 2010

Tabel B2: Registraties nationaliteit en leeftijd in 2010

Tabel B3: Registraties nationaliteit en woonregio in GBA in 2010

Tabel B4: Registraties nationaliteit en verblijfsduur in GBA in 2010

Tabel B5: Registraties nationaliteit en ouderschap in GBA in 2010

Tabel B6: Registraties nationaliteit en burgerlijke staat in GBA in 2010

C. Overzichtstabellen omvangschattingen MOE-landers in 2010

Tabel C1: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2010

Tabel C2: Omvangschatting GBA en nationaliteit in 2010

Tabel C3: Omvangschatting geslacht en nationaliteit in 2010

Tabel C4: Omvangschatting leeftijd en nationaliteit in 2010

Tabel C5: Omvangschatting woonregio en nationaliteit in 2010

Tabel C6: Omvangschatting verblijfsduur (GBA variabele) en nationaliteit in 2010

Tabel C7: Omvangschatting ouderschap (GBA variabele) en nationaliteit in 2010

Tabel C8: Omvangschatting burgerlijke staat (GBA variabele) en nationaliteit in 2010

Tabel C9: Omvangschatting GBA, geslacht en nationaliteit in 2010

Tabel C10: Omvangschatting leeftijd, geslacht en nationaliteit in 2010

A. Aantal in GBA, HKS en WNB geregistreerde MOE-landers

In Appendix A gaat het om de *geregistreerde* MOE-landers, *zonder* bijinschattingen van het aantal *niet*-geregistreerde MOE-landers. Hierbij zijn GBA en HKS gekoppeld en GBA en WNB gekoppeld.

1. De linker tabel (WNB) geeft een overzicht van het aantal geregistreerde MOE-landers met een registratie in GBA en/of WNB op de peildatum (laatste vrijdag van september). De tabel is gebruikt om het aantal MOE-landers te bepalen dat op de peildatum was geregistreerd in WNB maar niet in GBA,
2. De middelste tabel (HKS) geeft een overzicht van het aantal geregistreerde MOE-landers met een registratie in GBA en/of HKS in de laatste helft van het jaar. Deze tabel is gebruikt om het aantal MOE-landers te bepalen dat in de laatste helft van het jaar was geregistreerd in GBA, (Merk op dat het gekoppelde bestand van GBA en WNB een ondertelling geeft van het aantal registraties in GBA, omdat daar alleen registraties op de peildatum in voorkomen)
3. De rechter tabel (WNB+GBA) geeft een optelling van de registraties in WNB maar niet in GBA uit de linker tabel, en van de registraties in GBA uit de middelste tabel. Dit zijn dus MOE-landers met een registratie in WNB op peildatum en/of een registratie in GBA in de laatste helft van het jaar.

Tabel A1: Aantal in GBA geregistreerde MOE-landers naar nationaliteit in 2009

nationaliteit	WNB			HKS			WNB + GBA	
	ja	nee	Totaal	ja	nee	Totaal		Totaal
Pools	GBA	ja	19.282	16.239	35.521	ja	374	39.862
		nee	75.676	0	75.676	nee	1.373	0
								1.373
		Totaal	94.958	16.239	111.197	Totaal	1.747	41.235
Roemeens	GBA	ja	1.962	4.298	6.260	ja	65	7.091
		nee	955	0	955	nee	646	0
								646
		Totaal	2.917	4.298	7.215	Totaal	711	7.802
Bulgaars	GBA	ja	1.281	9.147	10.428	ja	133	12.083
		nee	236	0	236	nee	224	0
								224
		Totaal	1.517	9.147	10.664	Totaal	357	12.440
overig MOE	GBA	ja	6.261	6.858	13.119	ja	141	15.074
		nee	7.792	0	7.792	nee	554	0
								554
		Totaal	14.053	6.858	20.911	Totaal	695	15.769
Totaal			113.445	36.542	149.987	Totaal	3.510	73.736
								77.246
								159.108

Tabel A2: Aantal in GBA geregistreerde MOE-landers naar nationaliteit in 2010

nationaliteit	WNB			HKS			WNB + GBA	
	ja	nee	Totaal	ja	nee	Totaal		Totaal
Pools	GBA	ja	25.009	17.532	42.541	ja	481	47.181
		nee	82.615	0	82.615	nee	1.662	0
								47.662
		Totaal	107.624	17.532	125.156		2.143	47.181
Roemeens	GBA	ja	2.180	5.044	7.224	ja	91	8.166
		nee	966	0	966	nee	700	0
								8.257
		Totaal	3.146	5.044	8.190		791	8.166
Bulgaars	GBA	ja	1.498	10.168	11.666	ja	160	13.390
		nee	223	0	223	nee	219	0
								13.550
		Totaal	1.721	10.168	11.889		369	13.390
overig MOE	GBA	ja	8.287	7.998	16.285	ja	158	18.567
		nee	9.962	0	9.962	nee	627	0
								18.725
		Totaal	18.249	7.998	26.247		785	18.567
		Totaal	130.740	40.742	171.482		4.098	87.304
								91.402
								181.960

B. Overzichtstabellen van MOE-landers met een registratie in GBA in 2010

Tabellen B zijn tabellen van MOE-landers in GBA. Deze informatie is in principe vrij voor handen bij het CBS maar heeft een plek gekregen in het rapport voor een vergelijking met Tabellen C, waarin ook bijschattingen zijn opgenomen. Voor de vier groepen MOE-landers afzonderlijk worden in Tabellen B uitsplitsingen gemaakt naar geslacht, leeftijd, woonregio, verblijfsduur, ouderschap en burgerlijke staat. De waarde van het presenteren van tabel B in deze publicatie ligt in de vergelijking met de cijfers uit tabel C.

Tabel B1: Registraties nationaliteit en geslacht in GBA in 2010

Geslacht	Pools (n=47.662)	Roemeens (n=8.257)	Bulgaars (n=13.550)	Overig MOE (n=18.725)	Totaal (n=88.194)
Man	50%	43%	49%	44%	48%
Vrouw	50%	57%	51%	56%	52%
Totaal	100%	100%	100%	100%	100%

Tabel B2: Registraties nationaliteit en leeftijd in GBA in 2010

Leeftijd	Pools (n=47.662)	Roemeens (n=8.257)	Bulgaars (n=13.550)	Overig MOE (n=18.725)	Totaal (n=88.194)
15-24	18%	27%	27%	25%	22%
25-34	48%	46%	36%	47%	46%
35-49	27%	23%	33%	24%	27%
50-64	7%	4%	5%	5%	6%
Totaal	100%	100%	100%	100%	100%

Tabel B3: Registraties nationaliteit en woonregio in GBA in 2010

Regio	Pools (n=47.662)	Roemeens (n=8.257)	Bulgaars (n=13.550)	Overig MOE (n=18.725)	Totaal (n=88.194)
Groningen	1%	3%	3%	3%	2%
Friesland	1%	3%	1%	2%	2%
Drenthe	1%	1%	0%	1%	1%
IJsselland	2%	1%	1%	2%	1%
Twente	2%	2%	1%	2%	2%
Noord- en Oost-Gelderland	2%	2%	1%	2%	2%
Gelderland-Midden	2%	2%	2%	3%	2%
Gelderland-Zuid	3%	2%	1%	2%	2%
Utrecht	4%	6%	6%	7%	5%
Noord-Holland Noord	4%	3%	2%	3%	4%
Zaanstreek-Waterland	1%	1%	3%	2%	1%
Kennemerland	4%	3%	4%	4%	4%
Amsterdam-Amstelland	7%	23%	17%	15%	12%
Gooi- en Vechtstreek	1%	1%	1%	2%	1%
Haaglanden	14%	10%	28%	13%	15%
Hollands Midden	7%	3%	2%	3%	5%
Rotterdam-Rijnmond	11%	9%	16%	12%	12%
Zuid-Holland-Zuid	2%	2%	1%	2%	2%
Zeeland	2%	2%	1%	1%	2%
Midden- en West Brabant	9%	8%	4%	5%	7%
Brabant-Noord	5%	1%	1%	2%	3%
Brabant-Zuid-Oost	7%	5%	2%	5%	6%
Limburg-Noord	4%	2%	1%	2%	3%
Limburg-Zuid	2%	4%	1%	3%	2%
Flevoland	2%	2%	1%	2%	2%
Totaal	100%	100%	100%	100%	100%

Tabel B4: Registraties nationaliteit en verblijfsduur in GBA in 2010

Verblijfsduur	Pools (n=47.662)	Roemeens (n=8.257)	Bulgaars (n=13.550)	Overig MOE (n=18.725)	Totaal (n=88.194)
0-4mnd	10%	13%	14%	14%	12%
5-12mnd	18%	18%	17%	21%	18%
1-2jr	20%	19%	21%	19%	20%
2-5jr	34%	33%	41%	24%	33%
>5jr'	18%	17%	7%	22%	17%
Totaal	100%	100%	100%	100%	100%

Tabel B5: Registraties nationaliteit en ouderschap in GBA in 2010

Ouderschap	Pools (n=47.662)	Roemeens (n=8.257)	Bulgaars (n=13.550)	Overig MOE (n=18.725)	Totaal (n=88.194)
kinderen	26%	19%	18%	18%	22%
geen kinderen	74%	81%	82%	82%	78%
Totaal	100%	100%	100%	100%	100%

Tabel B6: Registraties nationaliteit en burgerlijke staat in GBA in 2010

	Pools (n=47.662)	Roemeens (n=8.257)	Bulgaars (n=13.550)	Overig MOE (n=18.725)	Totaal (n=88.194)
Burgerlijke staat					
ongehuwd	62%	70%	80%	77%	69%
getrouwd/samenwonend	33%	26%	18%	19%	27%
Gescheiden/verweduwd	5%	4%	2%	4%	4%
Totaal	100%	100%	100%	100%	100%

C. Omvangschattingen MOE-landers in 2010

Tabellen C zijn tabellen van zowel *geregistreerde* als *niet-geregistreerde* MOE-landers, waarbij deze laatste groep is bijgeschat op basis van de koppeling tussen GBA en HKS. In tabellen 1 en 2 worden aantallen gegeven, en in latere tabellen procentuele uitsplitsingen naar registratie in GBA, geslacht, leeftijd, woonregio, verblijfsduur, ouderschap en burgerlijke staat, en de kruising tussen registratie in GBA en geslacht, en de kruising tussen leeftijd en geslacht.

Tabel C1: Schatting geregistreerd versus niet-geregistreerd naar nationaliteit in 2010

nationaliteit	Geregistreerd	%	Niet-geregistreerd	%	Populatieschatting	%	95% betrouwbaarheidsinterval
Pools	130.277	77%	39.482	23%	169.759	50%	(159.976 – 182.707)
Roemeens	9.223	13%	60.002	87%	69.225	20%	(62.693 – 77.692)
Bulgaars	13.773	36%	25.011	64%	38.784	11%	(34.374 – 44.234)
Overig MOE	28.687	45%	34.640	55%	63.327	19%	(58.856 – 69.116)
Totaal	181.960	53%	159.134	47%	341.094	100%	(322.244 – 366.057)

Tabel C2: Omvangschatting GBA en nationaliteit 2010

Registratie	Pools (n=169.759)		Roemeens (n=69.225)		Bulgaars (n=38.784)		Overig MOE (n=63.327)	Totaal (n=341.094)
GBA								
Ja	27%		11%		34%		29%	25%
Nee	73%		89%		66%		71%	75%
Totaal	100%		100%		100%		100%	100%

Tabel C3: Omvangschatting geslacht en nationaliteit in 2010

geslacht	Pools (n=169.759)	Roemeens (n=69.225)	Bulgaars (n=38.784)	Overig MOE (n=63.327)	Totaal (n=341.094)
man	69%	54%	38%	68%	62%
vrouw	31%	46%	62%	32%	38%
Totaal	100%	100%	100%	100%	100%

Tabel C4: Omvangschatting leeftijd en nationaliteit in 2010

leeftijd	Pools (n=169.759)	Roemeens (n=69.225)	Bulgaars (n=38.784)	Overig MOE (n=63.327)	Totaal (n=341.094)
15-24	24%	36%	35%	32%	29%
25-34	45%	41%	33%	42%	42%
35-49	23%	18%	27%	20%	22%
50-64	8%	4%	5%	6%	7%
Totaal	100%	100%	100%	100%	100%

Tabel C5: Omvangschatting woonregio en nationaliteit in 2010

Regio	Pools (n=169.759)	Roemeens (n=69.225)	Bulgaars (n=38.784)	Overig MOE (n=63.327)	Totaal (n=341.094)
Groningen	1%	3%	3%	3%	2%
Friesland	1%	3%	1%	2%	2%
Drenthe	1%	1%	0%	1%	1%
IJsselland	2%	2%	1%	2%	1%
Twente	2%	2%	1%	2%	2%
Noord- en Oost Gelderland	2%	2%	1%	2%	2%
Gelderland Midden	2%	2%	2%	3%	2%
Gelderland Zuid	3%	2%	1%	2%	2%
Utrecht	4%	6%	6%	7%	5%
Noord-Holland Noord	4%	3%	2%	3%	3%
Zaanstreek-Waterland	1%	1%	3%	2%	1%
Kennermerland	4%	3%	4%	4%	4%
Amsterdam-Amstelland	7%	23%	17%	15%	13%
Gooi & Vecht	1%	1%	1%	2%	1%
Haaglanden	14%	10%	28%	13%	14%
Hollands midden	7%	3%	2%	3%	5%
Rotterdam Rijnmond	11%	9%	16%	12%	11%
Zuid-Holland-Zuid	2%	2%	1%	2%	2%
Zeeland	2%	2%	1%	1%	2%
Midden- en West-Brabant	9%	8%	4%	5%	7%
Brabant-Noord	5%	1%	1%	2%	3%
Brabant-Zuid-Oost	7%	5%	2%	5%	6%
Limburg-Noord	4%	2%	1%	2%	3%
Limburg-Zuid	2%	4%	1%	3%	2%
Flevoland	2%	2%	1%	2%	2%
Totaal	100%	100%	100%	100%	100%

Tabel C6: Omvangschatting verblijfsduur (GBA variabele) en nationaliteit in 2010

verblijfsduur	Pools (n=169.759)	Roemeens (n=69.225)	Bulgaars (n=38.784)	Overig MOE (n=63.327)	Totaal (n=341.094)
0-4 mnd	14%	18%	17%	19%	16%
5-12 mnd	25%	27%	21%	29%	26%
1-2 jr	21%	19%	21%	20%	20%
2-5 jr	32%	31%	38%	22%	30%
> 5 jr	8%	6%	4%	10%	7%
Totaal	100%	100%	100%	100%	100%

Tabel C7: Omvangschatting ouderschap (GBA variabele) en nationaliteit in 2010

ouderschap	Pools (n=169.759)	Roemeens (n=69.225)	Bulgaars (n=38.784)	Overig MOE (n=63.327)	Totaal (n=341.094)
kinderen	21%	15%	18%	13%	18%
geen kinderen	79%	85%	82%	87%	82%
Totaal	100%	100%	100%	100%	100%

Tabel C8: Omvangschatting burgerlijke staat (GBA variabele) en nationaliteit in 2010

Burgerlijke staat	Pools (n=169.759)	Roemeens (n=69.225)	Bulgaars (n=38.784)	Overig MOE (n=63.327)	Totaal (n=341.094)
ongehuwd	67%	75%	81%	81%	73%
getrouwd/samenwonend	29%	22%	17%	16%	24%
Gescheiden/verweduwd	4%	3%	2%	3%	3%
Totaal	100%	100%	100%	100%	100%

Tabel C9: Omvangschatting GBA, geslacht en nationaliteit in 2010

GBA	geslacht	Pools (n=169.759)	Roemeens (n=69.225)	Bulgaars (n=38.784)	Overig MOE (n=63.327)	Totaal (n=341.094)
ja	man	49%	39%	49%	43%	47%
	vrouw	51%	61%	51%	57%	53%
Totaal		100%	100%	100%	100%	100%
nee	man	77%	56%	33%	78%	67%
	vrouw	23%	44%	67%	22%	33%
Totaal		100%	100%	100%	100%	100%

Tabel C10: Omvangschatting leeftijd, geslacht en nationaliteit in 2010

leeftijd	geslacht	Pools (n=169.759)	Roemeens (n=69.225)	Bulgaars (n=38.784)	Overig MOE (n=63.327)	Totaal (n=341.094)
15-24	man	67%	50%	36%	69%	59%
	vrouw	33%	50%	64%	31%	41%
Totaal		100%	100%	100%	100%	100%
25-34	man	68%	54%	37%	66%	62%
	vrouw	32%	46%	63%	34%	38%
Totaal		100%	100%	100%	100%	100%
35-49	man	72%	59%	42%	68%	65%
	vrouw	28%	41%	58%	32%	35%
Totaal		100%	100%	100%	100%	100%
50-64	man	71%	58%	39%	76%	67%
	vrouw	29%	42%	61%	24%	33%
Totaal		100%	100%	100%	100%	100%