



rijksuniversiteit
groningen

Wie trouwt met wie?

Een kwantitatief onderzoek naar de invloed van co-etnische sociale netwerken in de wijk op individuele partnerkeuze en verschillen hierin tussen etnische groepen en 1^e en 2^e generatie migranten

Naam: J.D. Slump (s3736482)

E-mail: j.d.slump@student.rug.nl

Opleiding: Sociologie (bachelor)

Faculteit: Gedrags- en maatschappijwetenschappen

Scriptiebegeleider: AP. J. Nieuwenhuis

Tweede Lezer: dr. R.C. Smaniotto

Bachelorscriptie

8 Juni 2022

Abstract

De jarenlange toestroom van migranten heeft ervoor gezorgd dat er een multiculturele samenleving is ontstaan in Nederland, waardoor de kans dat iemand een interetnisch huwelijk aangaat is toegenomen. Hoewel trouwen met een Nederlander als laatste stap kan worden gezien van volledige integratie van mensen met een migratieachtergrond, kiezen velen van hen om te trouwen met iemand die dezelfde etniciteit heeft (co-etnisch). Co-etnische sociale netwerken uit de wijk waar deze mensen wonen zouden hier invloed op kunnen hebben.

Het doel van dit onderzoek is om te achterhalen in hoeverre co-etnische sociale netwerken in de wijk invloed hebben op iemands keuze voor een co-etnische partner. Hierbij zal onderzocht worden in hoeverre de sterkte van dit verband verschilt voor verschillende etnische groepen en voor 1^e en 2^e generatie migranten. De etnische groepen die onderzocht zullen worden zijn de vier grootste etnische minderheidsgroepen in Nederland, namelijk mensen met een Turkse, Marokkaanse, Surinaamse en Antilliaanse migratieachtergrond.

Om dit te kunnen achterhalen is gebruik gemaakt van data afkomstig uit de Netherlands Longitudinal Lifecourse Study (NeLLS) wave 1 (2009). De dataset bevat een over-sample van Turkse en Marokkaanse Nederlanders en bestaat daarnaast uit twee waves, is grootschalig (>5000 respondenten in wave 1) en er zijn zowel kwantitatieve als kwalitatieve methoden ingezet om de data te verkrijgen.

De resultaten uit de logistische regressieanalyse ($n=953$) laten zien dat mensen die één van de eerdergenoemde migratieachtergronden heeft meer kans hebben om zich in een co-etnisch huwelijk te bevinden wanneer zij geregeld of wekelijks contact hebben met co-etnische buurtgenoten, dan wanneer zij vrijwel nooit contact hebben met co-etnische buurtgenoten. Echter, uit de resultaten is niet gebleken dat dit verband verschilt in sterkte tussen de etnische groepen of tussen 1^e en 2^e generatie migranten.

Voor vervolgonderzoek is het belangrijk om een representatieve steekproef te hebben en alternatieve verklaringen mee te nemen. Daarnaast zouden gemeenten meer kunnen investeren in buurthuizen en andere ontmoetingsplekken. Door mensen met een migratieachtergrond meer in contact te brengen met Nederlanders, zou de integratie van deze mensen sneller kunnen verlopen.

Inhoudsopgave

<i>Inleiding</i>	4
<i>H1: Theorie</i>	7
<i>H2: Methoden</i>	13
2.1 Dataset en steekproef	13
2.2 Onderzoeksdesign	14
2.3 Operationalisatie.....	15
2.4 Analyseplan	19
<i>H3: Resultaten</i>	20
3.1 Beschrijvende statistieken	20
3.2 Modevaluatie	23
3.2.1. Assumptie, multicollineariteit & outliers	24
3.3 Bespreking belangrijke resultaten en hypothesetoetsing	25
<i>H4: Conclusie en discussie</i>	30
Conclusie	30
Discussie.....	32
<i>Literatuurlijst</i>	35
<i>Bijlagen</i>	38
Bijlage 1: Gebruikte variabelen en hun bewerkingen.....	38
Bijlage 2: Bespreking statistieken en analyses	51
Bijlage 3: Assumptietoets, outliers en multicollineariteit.....	65

Inleiding

Migratie is van alle tijden; decennialang al komen mensen vanuit verschillende landen naar Nederland toe en het aandeel mensen met een migratieachtergrond¹ in Nederland blijft stijgen (CBS, 2020). De grootste groep mensen met een migratieachtergrond in Nederland zijn mensen van Turkse en Marokkaanse afkomst. Door de aanwezigheid van deze groepen mensen met een migratieachtergrond is in Nederland een multiculturele samenleving ontstaan die steeds diverser lijkt te worden. Dit zorgt op zijn beurt dat de kans groter wordt dat iemand een interetnisch huwelijk, dat wil zeggen een huwelijk tussen twee mensen die een verschillende etniciteit hebben, aan zal gaan. Het aangaan van een interetnisch huwelijk met een inheemse² kan worden gezien als de laatste stap die een persoon met een migratieachtergrond kan zetten om volledig te integreren (Tolsma et al., 2007).

Interetnische huwelijken zouden ervoor kunnen zorgen dat grenzen tussen etnische groepen vervagen en vooroordelen over andere etnische groepen verminderen (Hartung et al., 2011). Dergelijke huwelijken worden hierdoor ook gezien als een indicator van sociale cohesie binnen samenlevingen. Van alle voltrokken huwelijken in 2016 was 16% tussen iemand met een migratieachtergrond en een Nederlander³ (CBS, 2017).

Desondanks kiezen veel mensen met een migratieachtergrond ervoor om te trouwen met iemand die dezelfde etniciteit heeft (co-etnisch); slechts 20% van zowel de Turkse Nederlanders als de Marokkaanse Nederlanders heeft een echtgenoot(e) met een andere etniciteit (CBS, 2020). Daarentegen trouwden 50% van de Surinaamse Nederlanders interetnisch en 75% van de Antilliaanse Nederlanders (CBS, 2020). Deze cijfers laten zien dat iemands etnische achtergrond ook van invloed zou kunnen zijn op iemands keuze voor het al dan niet betreden van een interetnisch huwelijk.

¹ In dit onderzoek zal er op verschillende manieren verwezen worden naar mensen met een migratieachtergrond, zo zal er soms gesproken worden over bijvoorbeeld Turkse Nederlanders of over mensen met een Turkse migratieachtergrond. Deze voorbeelden kunnen gezien worden als synoniemen van elkaar. Daarnaast zal er gebruik worden gemaakt van de term "migranten" wanneer er gesproken wordt over 1^e en 2^e generatie migranten. Hoewel 2^e generatie migranten niet onder de officiële definitie van migranten vallen (zij zijn hier immers geboren), wordt deze groep doorgaans wel benoemd als 2^e generatie migranten.

² Ook wel 'native' of 'inboorling' genoemd, waarmee een persoon bedoeld wordt die behoort tot een land omdat diegene daar geboren is.

³ Wanneer er gesproken wordt over een 'Nederlander' betekent dit dat de ouders van deze persoon beiden in Nederland geboren zijn.

Zoals net is gesteld kan het aangaan van een interetnisch huwelijk met een inheemse een positief effect hebben op iemands integratieproces. Dit is ook te zien aan de Nederlandse statistieken; Surinaamse en Antilliaanse Nederlanders, die vaker een interetnisch huwelijk aangaan met bijvoorbeeld Nederlanders, hebben doorgaans minder moeite met de Nederlandse taal, voelen zich vaker thuis in Nederland, hebben vaker betaald werk, zijn vaker hoogopgeleid en zijn meer actief in het verenigingsleven dan Turkse en Marokkaanse Nederlanders (CBS Statline, 2011; CBS, 2022a; CBS, 2022b).

Naast iemands etnische achtergrond kunnen wijken ook invloed hebben op iemands keuze voor een co-etnisch of interetnisch huwelijk; gesegregeerde wijken waarin voornamelijk mensen wonen met een bepaalde migratieachtergrond kunnen huwelijken binnen dezelfde etnische groep faciliteren. Het nabijheidseffect en het principe van ruimtelijke homogamie zijn theorieën die hiermee in lijn zijn, deze theorieën suggereren namelijk dat mensen geneigd zijn om met mensen te trouwen die dicht bij hen wonen (Haandrikman, 2018). Het belang van ontmoetingsmogelijkheden wordt in dergelijke theorieën benadrukt en gezien als een facilitator van (interetnische) huwelijken (Carol, 2014; van Zantvliet et al., 2014; Haandrikman, 2018).

Er is al veel onderzoek gedaan naar wat voor mensen een co-etnisch of interetnisch huwelijk aangaan en welke factoren hierbij een rol kunnen spelen (onder andere Kalmijn, 1998; Kalmijn & van Tubergen, 2010; Carol, 2014). Zo blijkt dat bijvoorbeeld religie en opleidingsniveau een grote rol kunnen spelen in het wel of niet aangaan van een interetnisch huwelijk (Kalmijn, 1998). Daarnaast zijn er al verschillende theorieën gelinkt aan iemands partnerkeuze, zoals de eerdergenoemde twee.

Onderzoek doen naar de invloed van wijken op partnerkeuze sluit dan ook aan op deze theorieën, omdat mensen die in dezelfde wijk wonen nabij elkaar zijn en hierdoor meer ontmoetingsmogelijkheden hebben. Daarnaast kunnen co-etnische sociale netwerken binnen wijken een bepalende rol spelen in het vormen van iemands identiteit; iemands gevoel van 'thuis zijn' kan opgewekt worden door contact te hebben met co-etnische buurtgenoten waardoor diegene zich vervolgens zal identificeren met de wijk (van den Berg & Timmermans, 2015). Dit kan op zijn beurt de identificatie met co-etnische buurtgenoten weer versterken, wat invloed kan hebben op diens partner preferenties.

Er zijn echter weinig onderzoeken die op wijkniveau kijken naar wat specifiek de invloed is van de aanwezigheid van co-etnische sociale netwerken in de wijk, op de keuze

van iemand met een bepaalde migratieachtergrond om te trouwen met iemand die dezelfde etniciteit heeft en hierin onderscheid maken tussen verschillende etnische groepen en 1^e en 2^e generatie migranten. Dit is interessant om te onderzoeken, omdat het inzicht kan geven in de invloed van iemands leefomgeving op diens partnerkeuze en in hoeverre de rol van de omgeving hierin sterker of zwakker is voor bepaalde etnische groepen en voor 1^e en 2^e generatie migranten.

Omdat 2^e generatie migranten in Nederland veelal beter gesocialiseerd zijn en de Nederlandse taal beter spreken dan de 1^e generatie (Kalmijn & van Tubergen, 2007; Hartung et al., 2011), zijn co-etnische sociale netwerken uit hun wijk wellicht minder van invloed op hun partnerkeuze. Hetzelfde geldt voor Surinaamse en Antilliaanse Nederlanders; omdat zij doorgaans beter geïntegreerd lijken te zijn dan Turkse en Marokkaanse Nederlanders, zou het kunnen zijn dat dergelijke netwerken een minder grote rol spelen in hun partnerkeuze.

Dit leidt ons tot de onderzoeksvraag die centraal zal staan in dit onderzoek en luidt: *“Hangt een hoge mate aan co-etnische sociale netwerken in de wijk samen met individuele partnerkeuze en verschilt dit voor verschillende etnische groepen en voor 1^e en 2^e generatie migranten?”*

In dit onderzoek zal allereerst een overzicht worden gegeven van de literatuur die beschikbaar is over interetnische en co-etnische huwelijken en verschillen daarin tussen etnische groepen en 1^e en 2^e generatie migranten. Vervolgens zullen, op basis van de data afkomstig uit de NeLLS (Netherlands Longitudinal Lifecourse Study) dataset (Tolsma et al., 2014) bestaande uit een over-sample van Turkse en Marokkaanse Nederlanders, descriptieve resultaten over de al dan niet aanwezige co-etnische huwelijken aan bod komen. Een logistische regressie moet uitwijzen wat de invloed van co-etnische sociale netwerken in de wijk is op iemands partnerkeuze, waaruit tot slot conclusies getrokken kunnen worden.

H1: Theorie

Het Engelse gezegde “birds of a feather flock together” suggereert dat mensen die op elkaar lijken geneigd zijn met elkaar om te gaan. Ook bij het uitzoeken van een partner lijkt dit op te gaan; het concept endogamie stelt bijvoorbeeld dat mensen vaak trouwen met iemand die uit dezelfde sociale groep komt, bijvoorbeeld dezelfde etnische groep (Kalmijn, 1998). Het delen van dezelfde normen, waarden en overtuigingen zouden redenen hiervoor kunnen zijn. In dergelijke huwelijken is er veelal sprake van onderling begrip en bevestiging (Kalmijn, 1998) en is de kans op een scheiding kleiner dan wanneer er weinig tot geen gelijkenissen tussen de partners zijn (Carol, 2014).

Naast het hebben van gelijkenissen met een partner, zijn er andere mechanismen die beïnvloeden wie iemand als partner kiest. Nabijheid en de mogelijkheid om een partner te ontmoeten zijn bijvoorbeeld ook van invloed op iemands partnerkeuze (Carol, 2014; van Zantvliet et al., 2014). De kans dat iemand een relatie met een ander aangaat is groter wanneer diegene meer mogelijkheden heeft om de ander te ontmoeten (Kalmijn, 1998). Beiden in dezelfde wijk wonen zou bijvoorbeeld de kans om elkaar vaker tegen te komen kunnen vergroten. Daarnaast draagt de grootte van de groep waar iemand onderdeel van is ook bij aan de kans om iemand te ontmoeten die onderdeel is van dezelfde groep; hoe groter de groep hoe groter de kans om een groepsgeenoot te ontmoeten (Kalmijn, 1998; Kalmijn & van Tubergen, 2010; van Zantvliet et al., 2014). Omdat in Nederland de meerderheid van de inwoners van Nederlandse origine zijn, is de kans dat een Nederlander met een Nederlander trouwt groot. Maar dit kan per wijk verschillen; sommige wijken hebben bijvoorbeeld een andere etnische samenstelling dan andere wijken (Kalmijn & van Tubergen, 2010).

Mensen met een migratieachtergrond zijn namelijk geneigd om in wijken te gaan wonen waarin de meerderheid van de inwoners dezelfde etniciteit als hen heeft, omdat zij op deze manier hun levensstijl kunnen behouden en dit hen zekerheid geeft dat zij hulp van anderen met dezelfde etniciteit zullen krijgen mochten zij dat nodig hebben (Bolt & van Kempen, 2010). Hierdoor zijn er meer ontmoetingsmogelijkheden met mensen van dezelfde etniciteit en is de kans op endogamie groter (Kalmijn, 1998). Wanneer er bijvoorbeeld veel Marokkaanse sociale netwerken in een wijk zijn, dat wil zeggen veel Marokkaans-

Nederlandse buurtgenoten die onderling contact met elkaar hebben, is de kans om iemand met een Marokkaanse achtergrond te ontmoeten relatief groot.

Kalmijn (1998) omschrijft wijken dan ook als 'lokale huwelijksmarkten' waarin buurtgenoten veelal dezelfde sociaaleconomische status, religie en etniciteit hebben maar wel verschillend zijn in sekse. Dit maakt dat het een ideale plek is om een partner te treffen. Maar zelfs als iemand geen partner uit de wijk heeft, kunnen sociale netwerken uit de wijk invloed hebben op iemands partnerkeuze. Sociale netwerken binnen wijken kunnen namelijk een bepalende rol spelen in het vormen van iemands identiteit; iemands gevoel van 'thuis zijn' kan opgewekt worden door contact te hebben met co-etnische buurtgenoten waardoor diegene zich vervolgens zal identificeren met de wijk (van den Berg & Timmermans, 2015). Dit kan op zijn beurt weer iemands gevoel van verbondenheid met co-etnische buurtgenoten versterken, door een gedeelde identificatie met de wijk (Fong et al., 2019), en uiteindelijk invloed hebben op iemands partner preferenties. Deze gedeelde identificatie zou bijvoorbeeld versterkt kunnen worden door lokaal beleid waar iedereen in de wijk zich aan moet houden en de aanwezigheid van lokale politieke partijen. Dit maakt dat co-etnische sociale netwerken in wijken een unieke rol kunnen spelen in iemands partnerkeuze, een rol die anders is dan de rol van dergelijke netwerken op school of op werk.

Daar komt bij dat, volgens het nabijheidseffect, mensen sneller bevriend raken met iemand die dicht bij hen woont (Festinger, 1973). Herhaaldelijk blootgesteld worden aan iemand, bijvoorbeeld mensen dagelijks zien, kan ertoe leiden dat je meer vertrouwd met de betreffende persoon raakt wat op zijn beurt weer voor een aantrekkingskracht kan zorgen (Batoool & Malik, 2010). Mensen houden namelijk van mensen en objecten die hen bekend voorkomen (Keizer, 2021). Het principe van ruimtelijke homogamie suggereert hetzelfde; mensen zijn geneigd met iemand te trouwen die dicht bij hen woont (Haandrikman, 2018). Onderzoek van Haandrikman (2018) laat zien dat wanneer de afstand toeneemt tussen mogelijke partners, de kans op een match afneemt. Dit leidt ons tot de eerste hypothese:

"Hoe hoger de mate van co-etnische sociale netwerken in de wijk, des te groter de kans dat iemand uit die wijk kiest voor een partner die dezelfde etniciteit heeft." (H1)

Maar geldt dit voor alle mensen met een migratieachtergrond of is de kans op endogamie kleiner voor bepaalde mensen met een migratieachtergrond, ongeacht in wat voor wijk zij wonen? Ik verwacht dat wanneer er een hoge mate co-etnische netwerken in de wijk zijn, 1^e generatie migranten sneller geneigd zijn met iemand te trouwen die dezelfde etnische achtergrond heeft dan 2^e generatie migranten. De 2^e generatie trouwt vaker interetnisch dan de 1^e generatie, omdat zij veelal beter gesocialiseerd en geïntegreerd zijn (Kalmijn & van Tubergen, 2007; Hartung et al., 2011). Zij spreken de dominante taal al, kennen de cultuur beter en hebben veelal meer ontmoetingsmogelijkheden met mensen van andere ethniciteiten (Hartung et al., 2011). Deze ontmoetingsmogelijkheden ontstaan doordat zij bijvoorbeeld naar een Nederlandse school gaan of doordat zij, omdat zij over het algemeen beter opgeleid zijn dan de 1^e generatie (Kalmijn & van Tubergen, 2007; SCP, 2020), op werk meer interetnisch contact hebben. Deze ontmoetingsmogelijkheden kunnen zich dus buiten de wijk om bevinden. Dit leidt ons tot de tweede hypothese:

“Het positieve effect van de mate van co-etnische sociale netwerken in de wijk op de kans dat iemand uit die wijk kiest voor een partner die dezelfde ethniciteit heeft is sterker voor 1^e generatie migranten dan voor 2^e generatie migranten.” (H2)

Maar is er ook nog verschil tussen etnische groepen, wanneer gekeken wordt naar de invloed van co-etnische sociale netwerken in wijken op iemands keuze voor een co-etnisch huwelijk? Ik verwacht dat co-etnische sociale netwerken in de buurt meer invloed hebben op de keuze van Turkse en Marokkaanse Nederlanders voor een co-etnisch huwelijk dan op de keuze van Surinaamse en Antilliaanse Nederlanders. Zoals eerder genoemd lijken Surinaamse en Antilliaanse Nederlanders namelijk doorgaans beter geïntegreerd te zijn dan Turkse en Marokkaanse Nederlanders. Zo spreken Surinaamse en Antilliaanse Nederlanders de Nederlandse taal veelal beter, voelen zij zich vaker thuis in Nederland en zijn zij vaker hoogopgeleid dan Turkse en Marokkaanse Nederlanders (CBS, 2022b). Daarnaast hebben zij ook vaker betaald werk en zijn meer actief in het verenigingsleven dan Turkse en Marokkaanse Nederlanders (CBS Statline, 2011; CBS, 2022a). Zo zijn 31% van de Turkse Nederlanders en 27% van de Marokkaanse Nederlanders in 2021 hoogopgeleid ten opzichte van 44% van de Surinaamse Nederlanders en 42% van de Antilliaanse Nederlanders (CBS, 2022b). Daarbij komt dat 62% van de Turkse Nederlanders en 61% van de Marokkaanse

Nederlanders aangaf in 2011 zich thuis te voelen in Nederland ten opzichte van 73% van de Surinaamse Nederlanders en 72% van de Antilliaanse Nederlanders (CBS Statline, 2011). 23% Van de Turkse Nederlanders en 16% van de Marokkaanse Nederlanders gaven toen ook aan moeite te hebben met het voeren van een gesprek in het Nederlands ten opzichte van 1% van de Surinaamse Nederlanders en 3% van de Antilliaanse Nederlanders (CBS Statline, 2011). Surinaamse en Antilliaanse Nederlanders lijken dan ook meer ontmoetingsmogelijkheden te hebben met bijvoorbeeld Nederlanders op school of op werk. Daarnaast voelen zij zich meer thuis in Nederland, wat wellicht kan komen doordat zij meer ontmoetingsmogelijkheden hebben met Nederlanders en hierdoor meer opgenomen zijn in de maatschappij. Dit kan op zijn beurt ervoor zorgen dat de co-etnische sociale netwerken in hun wijk een minder bepalende rol spelen in hun partnerkeuze, wat ons leidt tot de derde hypothese:

“Het positieve effect van de mate van co-etnische sociale netwerken in de wijk op de kans dat iemand uit die wijk kiest voor een partner die dezelfde etniciteit heeft is sterker voor Turkse en Marokkaanse Nederlanders dan voor Surinaamse en Antilliaanse Nederlanders.”
(H3)

Naast nabijheid, ontmoetingsmogelijkheden en gelijkenissen zijn er echter nog andere mechanismen die van invloed kunnen zijn op iemands keuze voor een partner die dezelfde etniciteit heeft, zoals iemands sekse. Mannen met een migratieachtergrond trouwen sneller interetnisch dan vrouwen met een migratieachtergrond, omdat vrouwen doorgaans voor de kinderen zorgen en daardoor meer druk ervaren van mensen van buitenaf, zoals familie (Hartung et al., 2011). Daarnaast zijn het binnen de moslimgemeenschap vaker de vrouwen die religie overbrengen op de kinderen; wanneer een moslimvrouw trouwt met iemand die geen moslim is, dan zouden hun kinderen verloren zijn voor de islam (Kalmijn & van Tubergen, 2007). Dit maakt dat vrouwen met een migratieachtergrond wellicht minder snel interetnisch trouwen dan mannen met een migratieachtergrond.

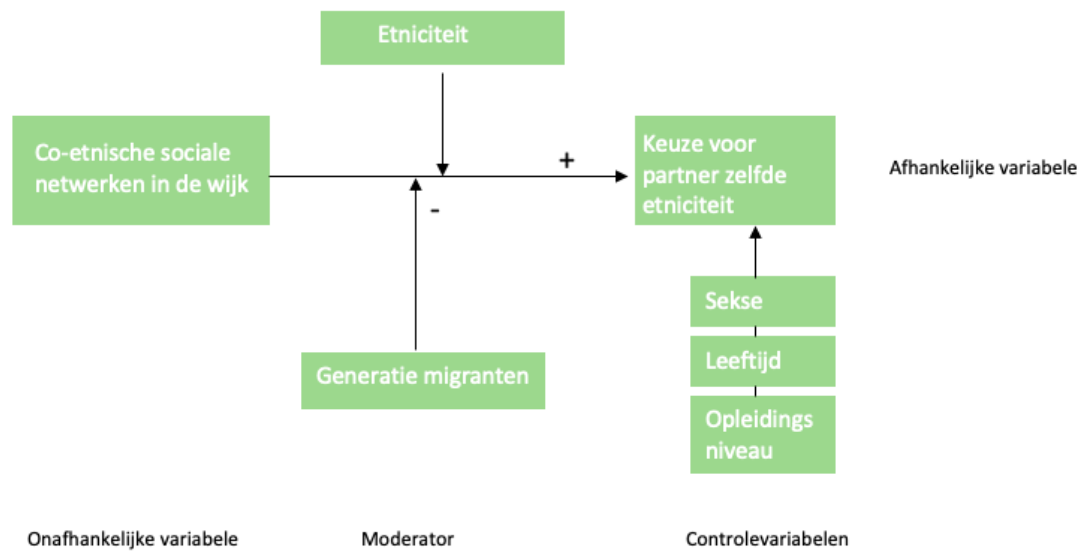
Daarnaast kan leeftijd ook van invloed zijn op iemands keuze voor een partner die dezelfde etniciteit heeft. De kans is groter dat jongere mensen met een migratieachtergrond interetnisch trouwen dan ouderen, omdat zij minder gesocialiseerd zijn in hun land van herkomst; zij doorgaans meer contact hebben met mensen van andere etnische groepen

doordat zij bijvoorbeeld naar school gaan; zij de dominante taal beter spreken (Kalmijn & van Tubergen, 2007). De kans dat bijvoorbeeld een Marokkaanse Nederlander trouwt met een Nederlander is kleiner wanneer de Marokkaanse Nederlander niet goed Nederlands spreekt (Hartung et al., 2011).

Opleidingsniveau kan daarnaast ook een rol spelen in iemands keuze voor een partner die dezelfde etniciteit heeft. De kans dat hoger opgeleide mensen met een migratieachtergrond interetnisch trouwen is groter dan de kans dat lager opgeleide mensen met een migratieachtergrond dat doen, omdat hoger opgeleiden meer ontmoetingsmogelijkheden hebben met bijvoorbeeld Nederlanders en hoger opgeleiden doorgaans een minder traditionele denkwijze hebben (van Zantvliet et al., 2014). Zij zouden dus meer open staan voor een interetnisch huwelijk en vaker Nederlanders ontmoeten, doordat zij bijvoorbeeld naar universiteiten gaan waar de meerderheid van de studenten Nederlands zijn (Kalmijn & van Tubergen, 2007).

Tot slot kan religie ook van invloed zijn op iemands keuze voor een partner die dezelfde etniciteit heeft en speelt hier doorgaans zelfs een grote rol in (Kalmijn & van Tubergen, 2010). Mensen die verbonden zijn aan traditionele religies en religies waarbinnen de kerk een grote rol speelt, zoals het katholicisme en islam, trouwen vaker co-etnisch dan mensen die daar niet aan verbonden zijn (Kalmijn, 1998). De sociale afstand tussen mensen die aan verschillende religies verbonden zijn is namelijk groter en bepaalde belangrijke waarden en normen worden wellicht niet gedeeld (Kalmijn & van Tubergen, 2010). Dit is ook terug te zien in de eerdergenoemde Nederlandse statistieken; de overwegend christelijke Surinaamse en Antilliaanse Nederlanders trouwen vaker interetnisch dan de overwegend islamitische Turkse en Marokkaanse Nederlanders. Maar, vanwege gebrek aan betrouwbare data over religie zal dit niet meegenomen worden in het verdere onderzoek.

De andere drie hierboven genoemde mechanismen zullen als controlevariabelen meegenomen worden in dit onderzoek en zijn tevens terug te vinden in Figuur 1. In Figuur 1 is het conceptueel model op een schematische wijze weergegeven, waaruit de hypothesen af te leiden zijn.



Figuur 1: Conceptueel model

H2: Methoden

2.1 Dataset en steekproef

De data die in dit onderzoek gebruikt zullen worden om de hypothesen te toetsen is afkomstig uit de NeLLS (Netherlands Longitudinal Lifecourse Study) dataset wave 1 (2009) (Tolsma et al., 2014). De dataset bevat een over-sample van de twee grootste etnische minderheidsgroepen in Nederland, namelijk Turkse en Marokkaanse Nederlanders. Daarnaast is de studie longitudinaal (panelonderzoek, 2 waves), grootschalig (>5000 respondenten in wave 1) en zijn er zowel kwantitatieve als kwalitatieve methoden ingezet (mixed methods). De dataset fungeert als bron van informatie van waaruit sociologen onderzoeken over normen en waarden, ongelijkheid en sociale cohesie kunnen opzetten. Omdat alleen gebruik gemaakt zal worden van de data afkomstig uit wave 1 zal informatie over wave 2 buiten beschouwing worden gelaten.

Een gestratificeerde steekproef is ingezet om representatieve data te verkrijgen, bestaande uit twee fasen (Tolsma et al., 2014). In de eerste fase zijn 35 gemeenten op quasi-willekeurige wijze geselecteerd door de gemeenten onder te verdelen per regio en urbanisatiegraad. Omdat het percentage Turkse en Marokkaanse Nederlanders erg laag is in kleine gemeenten op het platteland zijn deze gebieden niet meegenomen. Daarentegen moesten de vier grote westelijke steden Amsterdam, Rotterdam, Utrecht en Den Haag wel meegenomen worden om een representatieve steekproef van Turkse en Marokkaanse Nederlanders te verkrijgen. In de tweede fase zijn, door gemeenten zelf, respondenten op quasi-willekeurige wijze uit het geboorteregister geselecteerd op basis van etniciteit en leeftijd. Ook deze data bevat een over-sample van Turkse en Marokkaanse Nederlanders. De onderzoekers hebben aan de gemeenten doorgegeven hoeveel mensen geselecteerd moesten worden, waarna de gemeenten hen informatie gaf over de geselecteerde mensen hun naam, geboortedatum, geslacht en etniciteit.

Daarnaast is de dataset aangevuld met 39 statistieken over wijkgegevens, zoals de etnische samenstelling van wijken en gemiddelde leeftijd in wijken. Deze gegevens zijn afkomstig van het Centraal Bureau van de Statistiek (CBS).

Het responspercentage van wave 1 is 52% ($n=5312$), wat over het algemeen een gemiddeld respons is bij face-to-face enquêtes in Nederland (Tolsma et al., 2014). Interviewers gaven aan moeite te hebben om mensen aan de deur te krijgen, met name bij

de mensen met een migratieachtergrond. Het percentage Turkse en Marokkaanse Nederlanders die meewerkten was dan ook kleiner dan het percentage Nederlanders. De non-respons werd onder andere veroorzaakt doordat mensen niet bereikt konden worden of zij niet mee wilden werken. Mensen die verhuisd waren ten tijde van het veldwerk, die niet goed Nederlands spraken of die door andere onvoorziene omstandigheden niet mee konden werken zijn niet meegenomen in dit percentage.

De populatie waar mijn onderzoek op gericht is zijn getrouwde mensen die wonen in Nederland en een Turkse, Marokkaanse, Surinaamse of Antilliaanse migratieachtergrond hebben. De NeLLS dataset bevat onder andere informatie over de etniciteit en generatie van de respondenten, of zij getrouwd zijn of niet, wat de etniciteit is van hun partner, in hoeverre zij co-etnisch contact hebben met buurtgenoten en demografische kenmerken zoals leeftijd, sekse en opleidingsniveau. Omdat alle te meten concepten zijn opgenomen in de dataset, is het een geschikte informatiebron. Het is echter wel noodzakelijk om een selectie te maken van de steekproef, aangezien het onderzoek gaat over een specifieke doelgroep en de dataset ook informatie bevat over bijvoorbeeld Nederlandse en ongehuwde respondenten. Deze respondenten zullen dan ook niet meegenomen worden in de analyse. Daarnaast zullen ook de respondenten die een andere migratieachtergrond hebben dan Turks, Marokkaans, Surinaams of Antilliaans niet meegenomen worden in de analyse. Op deze manier blijven de respondenten behorend bij één van de grootste etnische minderheidsgroepen van Nederland nog over.

2.2 Onderzoeksdesign

Voor de eerste wave van het panelonderzoek zijn face-to-face interviews bij respondenten thuis afgenomen en is een enquête afgegeven die respondenten zelf moesten invullen. Het interview was voornamelijk gericht op sociaal-demografische en sociaal-economische kenmerken van de respondenten en de enquête op houdingen, sociale integratie en normen en waarden van de respondenten. De meeste interviews in wave 1 duurden ongeveer 45 minuten.

Voorafgaand aan het interview kregen de respondenten een informatiebrief en namen de interviewers persoonlijk contact met hen op, om zo ook zekerder te kunnen zijn van respons. Om het respons verder te vergroten werden respondenten 10,- aangeboden

wanneer zij mee zouden werken aan het interview, hopen dat dit hen zou motiveren ook de enquête in te vullen. Na afloop van het interview liet de interviewer de enquête achter bij de respondenten, die ook online ingevuld kon worden. Dit resulteerde echter in veel non-respons op de enquête. Halverwege deze fase van wave 1 hebben de onderzoekers besloten eerst aan de respondenten te vragen de enquête in te vullen voorafgaand aan het interview. Mochten respondenten dat niet gedaan hebben, dan zou de interviewer na afloop van het interview wachten totdat de respondenten alsnog de enquête ingevuld hadden.

Tijdens de face-to-face interviews zijn vragen gesteld die onderverdeeld zijn in tien verschillende gebieden, onder andere over het netwerk van de respondent, diens persoonlijkheid en wensen. Tweehonderddrie vragen zijn gesteld, soms bestaande uit korte subvragen of stellingen. De meeste vragen zijn multiple choice en sommigen kunnen alleen beantwoord worden met ja of nee. In de enquête, bestaande uit 97 vragen, worden negen verschillende gebieden bevraagd die deels overlappen met de gebieden uit het interview.

2.3 Operationalisatie

De concepten uit het conceptueel model zullen gemeten worden aan de hand van de variabelen die in de NeLLS dataset wave 1 zijn opgenomen. Hier is voor gekozen omdat wave 1 het grootste respons heeft; sommige respondenten zijn na wave 1 uitgevallen.

Zo zal de onafhankelijke variabele *co-etnische sociale netwerken in de wijk* gemeten worden aan de hand van meerdere variabelen, namelijk de etniciteit van de respondent, het geboorteland van de respondent en diens ouders en het contact dat de respondent heeft met co-etnische buurtgenoten. De zelfgerapporteerde etniciteit van de respondent heeft de volgende antwoordopties: (1) Marokkaanse 1^e generatie, (2) Marokkaanse 2^e generatie, (3) Turkse 1^e generatie, (4) Turkse 2^e generatie, (5) niet-westerse 1^e generatie, (6) niet-westerse 2^e generatie, (7) westerse 1^e generatie, (8) westerse 2^e generatie, en (9) Nederlandse. Omdat de generatie waar de respondent toe behoort voor het meten van deze variabele niet uitmaakt, zullen de 1^e en 2^e generatie samengevoegd worden. Tevens zullen de respondenten met een westerse en Nederlandse achtergrond uit de dataset gehaald worden, waardoor alleen de 1^e en 2^e generatie Marokkanen, Turken en niet-westerse mensen over zullen blijven. Omdat de zelfgerapporteerde etniciteit geen aparte Surinaamse

en Antilliaanse categorie bevat, zal er voor deze groepen respondenten gekeken worden naar hun geboorteland en die van hun ouders. Op deze manier kunnen 1^e en 2^e generatie Surinaamse en Antilliaanse respondenten geïdentificeerd worden. De bijbehorende vraag die aan de respondenten gesteld is luidt: "In welk land zijn de volgende personen geboren? U zelf/uw vader/uw moeder." Met als antwoordopties: (1) Nederland, (2) Marokko, (3) Turkije, (4) Suriname, (5) Nederlandse Antillen en (6) Ander land, namelijk... Respondenten die zelf geboren zijn in Suriname of de Antillen zullen bestempeld worden als 1^e generatie Surinaams/Antilliaans en respondenten die zelf niet geboren zijn in Suriname/de Antillen maar waarvan in ieder geval één van de ouders wel, zullen bestempeld worden als 2^e generatie Surinaams/Antilliaans (zie bijlage 1 voor de totstandkoming). De respondenten die een andere niet-westerse migratieachtergrond hebben zullen uit de dataset gehaald worden. De variabele die informatie bevat over de Turkse en Marokkaanse respondenten (zelfgerapporteerde etniciteit) zal samengevoegd worden met de groepen Surinaamse en Antilliaanse respondenten. Surinaamse en Antilliaanse respondenten zijn hierin ondergebracht in één groep, omdat beide groepen erg klein zijn (zie bijlage 1 tabel 2). Deze nieuwe variabele bevat dan drie antwoordopties: 0) Marokkaans, 1) Turks en 2) Antilliaans/Surinaams.

Daarnaast zal het contact dat respondenten hebben met co-etnische buurtgenoten meegenomen worden. De vraag uit de vragenlijst behorend bij deze variabele luidt: "Hoe vaak heeft u in uw buurt persoonlijk contact met iemand met de volgende etnische herkomst?" Per etnische groep (Nederlands, Turks, Marokkaans, Surinaams/Antilliaans, andere niet-westerse herkomst) zijn de antwoordopties: (1) (bijna) elke dag, (2) een of meerdere keren per week, (3) een aantal keer per maand, (4) ongeveer één keer per maand, (5) een aantal keren per jaar, (6) ongeveer één keer per jaar, (7) nooit en (8) niet van toepassing (heb deze persoon niet). Omdat alleen co-etnisch contact gemeten dient te worden, zullen de items die gaan over Nederlandse en andere niet-westerse buurtgenoten niet meegenomen worden. Van de overige drie items zullen allereerst de schalen gespiegeld worden, zodat een hogere waarde intensiever contact betekent. Omdat geen enkele respondent een 8 scoorde op de items, is deze optie achterwege gelaten (zie bijlage 1 figuren 1, 2 en 3). Daarna zullen een aantal antwoordopties samengevoegd worden, omdat er maar weinig respondenten zijn die vrijwel nooit contact hebben met co-etnische buurtgenoten (zie bijlage 1 tabel 3). Opties 6 en 7 zullen samengenomen worden waarbij de

nieuwe score 0 is ("vrijwel nooit"), 3, 4 en 5 samen krijgen een nieuwe score van 1 ("geregeld maar <1x per week") en opties 1 en 2 de score 2 ("minstens 1x per week"). Vervolgens zal de etniciteit van de respondenten vergeleken worden met de variabele die aangeeft hoeveel contact de respondenten hebben met co-etnische buurtgenoten, waarna de drie ontstane variabelen die per etnische groep co-etnisch contact weergeven op hun beurt samengevoegd zullen worden. Hierdoor ontstaat er één variabele die aangeeft hoe intensief dit co-etnisch contact is (zie bijlage 1 voor de totstandkoming). Co-etnische sociale netwerken in de wijk wordt dus gemeten door te kijken naar de intensiteit van het contact dat respondenten hebben met co-etnische buurtgenoten (kwaliteit) en niet de hoeveelheid contacten die zij hebben (kwantiteit).

De afhankelijke variabele *keuze voor partner zelfde etniciteit* is ook gemeten aan de hand van meerdere variabelen, namelijk de etniciteit van de respondent en het geboorteland van diens partner en de ouders van de partner. De antwoordopties bij de vraag in welk land de partner, diens vader en diens moeder is geboren zijn: (1) Nederland, (2) Marokko, (3) Turkije, (4) Suriname, (5) Nederlandse Antillen, en (6) Ander land, namelijk.... Allereerst zal de etniciteit van de partner vastgesteld worden aan de hand van het geboorteland van de partner en diens ouders (zie bijlage 1 voor de totstandkoming). De definitie van etniciteit die gehanteerd zal worden is de definitie die gebruikt is in de NeLLS studie en luidt: "De exacte herkomst wordt bepaald door het geboorteland van de moeder als beide ouders (of alleen de moeder) buiten Nederland zijn geboren. Als de moeder in Nederland is geboren, wordt de herkomst bepaald door het geboorteland van de vader. Een persoon is van Nederlandse afkomst als beide ouders in Nederland zijn geboren, ongeacht het eigen geboorteland (Tolsma et al., 2014)." Het zou dus kunnen dat de partner een 2^e generatie migrant is, maar de 1^e en 2^e generatie zijn samengenomen aangezien in dit geval alleen de etniciteit ertoe doet. Na het vaststellen van de etniciteit van de partner zal de schaal gelijkgesteld worden aan de schaal van de respondent's etniciteit. Dan is het mogelijk de etniciteit van beiden te vergelijken en te achterhalen welke respondenten een interetnisch huwelijk aangegaan zijn (score nieuw variabele is 0) en welke een co-etnisch huwelijk (score 1).

De variabele *generatie migranten* is gemeten aan de hand van de zelfgerapporteerde etniciteit van de respondenten (voor Turken en Marokkanen) en de eerder gemaakte variabelen die aangeven welke respondenten 1^e generatie Surinaams/Antilliaans zijn en

welke respondenten 2^e generatie Surinaams/Antilliaans. Omdat nu alleen de generatie van belang is en niet de etniciteit, zijn alle 1^e generatie migranten samengevoegd (score 1) en alle 2^e generatie migranten (score 0). Deze variabele is tevens onderdeel van een interactieterm en zal hiervoor vermenigvuldigd worden met *co-etnische sociale netwerken in de wijk*.

De etniciteit variabele die gebruikt is voor het meten van co-etnische sociale netwerken zal ook gebruikt worden voor de nieuwe variabele *eticiteit*, waarin er geen onderscheid gemaakt zal worden tussen de generaties. Maar, omdat deze variabele onderdeel zal zijn van een regressieanalyse moet deze gedummificeerd worden. Er zullen twee dummy's gemaakt worden; één die aangeeft of iemand Marokkaans is (1) of niet (0) en één die aangeeft of iemand Turks is (1) of niet (0). Wanneer iemand een Surinaamse of Antilliaanse migratieachtergrond heeft zal diegene op beiden een 0 scoren. Deze dummy's zijn ook onderdeel van twee interactietermen en beiden zullen hiervoor los van elkaar vermenigvuldigd worden met *co-etnische sociale netwerken in de wijk*.

Tot slot zullen ook een aantal controlevariabelen gebruikt worden, waaronder *seks* bestaande uit de opties (1) man en (2) vrouw. Deze zal tot een dummy gehercodeerd worden zodat *seks* opgenomen kan worden in de analyse. Mannen scoren dan een 0 en vrouwen een 1. *Leeftijd* is gemeten door te kijken naar de leeftijd van de respondenten ten tijde van het interview en varieert van 15 tot 49 jaar. Tot slot zal *opleidingsniveau* opgenomen worden in de analyse, die gemeten is aan de hand van verschillende items. Veertien verschillende opleidingen/niveaus zijn bevroegd en respondenten moesten aangeven of zij deze opleidingen hebben afgerond met een diploma (ja = 1, nee = 2). Van deze items is een continue variabele gemaakt die aangeeft wat de respondenten hun hoogst behaalde diploma is, variërend van alleen lagere school (score 1) tot aan promotietraject (score 11). Drie items zijn ondergebracht onder dezelfde waarden als drie andere items, namelijk "buitenlandse opleiding, niet goed in te delen, lager onderwijs", "buitenlandse opleiding, niet goed in te delen, middelbaar onderwijs" en "buitenlandse opleiding, niet goed in te delen, hoger onderwijs." Deze zijn ondergebracht onder dezelfde nieuwe waarden van respectievelijk "lagere school" (1), "mbo-kort..." (6) en "hbo" (8). Zie bijlage 1 voor de totstandkoming en een uitgebreide toelichting.

2.4 Analyseplan

Voor dit onderzoek zal door middel van een binaire logistische regressie modellen geschat worden. Zowel de afhankelijke als de onafhankelijke variabelen zijn namelijk dichotome variabelen. Model 1 zal de afhankelijke variabele *keuze voor partner zelfde etniciteit* bevatten en de controlevariabelen *sekse, leeftijd, religie* en *opleidingsniveau*. Op deze manier kan gecontroleerd worden wat voor invloed deze variabelen hebben op het al dan niet betreden van een co-etnisch huwelijk. Vervolgens zal in model 2 de onafhankelijke variabele *co-etnische sociale netwerken in de wijk* toegevoegd worden. Dan zal duidelijk worden of de resultaten in lijn zijn met hypothese 1; wat de invloed is van het al dan niet hebben van contact met co-etnische buurtgenoten op iemands keuze voor een co-etnisch huwelijk, gecontroleerd voor de controlevariabelen. Vervolgens zal in model 3 de dummy variabele *generatie* toegevoegd worden, om te meten wat de invloed van iemands generatie is op het aangaan van een co-etnisch huwelijk. Tevens zal dit model de twee dummy's behorende bij *etniciteit* bevatten. In model 4 zal de interactieterm toegevoegd worden van *generatie* en *co-etnische sociale netwerken in de wijk* om zo te kunnen controleren of de sterkte van het hoofdverband verschillend is voor 1^e en 2^e generatie migranten. Dit model zal hypothese 2 toetsen. Model 5 zal ook de interactieterm van de *etniciteit* dummy's en *co-etnische sociale netwerken in de wijk* bevatten, om zo ook verschillen tussen etnische groepen zichtbaar te krijgen wat betreft het hoofdverband. Dit laatste model zal hypothese 3 toetsen.

H3: Resultaten

In deze sectie zullen de resultaten besproken worden die uit de binaire logistische regressieanalyse zijn gekomen. Allereerst zullen beschrijvende statistieken besproken worden, die laten zien hoe de variabelen die onderdeel zijn van de analyse verdeeld zijn en in hoeverre zij met elkaar samenhangen. Daarna zullen de geschatte modellen geëvalueerd worden, waarbij gekeken wordt welk model het beste bij de data past aan de hand van verschillende statistische toetsen. Er zal hierbij ook besproken worden in hoeverre er voldaan is aan de assumptie behorend bij logistische regressie, in hoeverre er sprake is van multicollineariteit en of de data outliers bevat. Tot slot zullen de hypothesen getoetst worden aan de hand van de resultaten uit de best passende modellen.

3.1 Beschrijvende statistieken

In tabel 1 zijn de univariate statistieken weergegeven die de verdelingen van de variabelen laten zien die onderdeel zijn van de analyse. Opvallend is dat veel variabelen een enigszins scheve verdeling hebben. Zo zijn de meeste respondenten van de 1^e generatie (82.6%), hebben respondenten wekelijks contact met co-etnische buurtgenoten (72.2%), hebben zij vrijwel allemaal een partner die dezelfde etniciteit heeft (94.3%) en zijn vrijwel alle respondenten van Turkse of Marokkaanse komaf (respectievelijk 50.2% en 47.6%). Dit zou een negatieve invloed kunnen hebben op de generaliseerbaarheid van dit onderzoek, maar daar zal in de discussiesectie verder op ingegaan worden (zie pagina X). Verder is het opvallend dat er geen respondenten zijn die ouder dan 49 jaar zijn. 50+ers zijn dus geen onderdeel van de analyse, waardoor de resultaten wellicht niet generaliseerbaar zijn naar deze groep (zie discussiesectie).

Tabel 1: frequentieverdelingen van de in de analyse opgenomen variabelen: absolute aantallen en percentages (N=1254)

Variabele		Absolute aantallen^a	Percentages^b
1. Leeftijd		35 (6)	15 - 49
2. Sekse	0 = man	585	46.7%
	1 = vrouw	669	53.5%
3. Opleidingsniveau ^c (schaal 11 items)		4.7 (2.9)	1 - 11
4. Generatie migranten ^d	0 = 2 ^e generatie	217	17.4%
	1 = 1 ^e generatie	1033	82.6%
5. Co-etnische sociale netwerken in de wijk ^e	0 = vrijwel nooit	117	10.7%
	1 = geregeld maar <1x per week contact	187	17.1%
	2 = minstens 1x per week	788	72.2%
6. Zelfde etniciteit respondent en partner	0 = niet	71	5.7%
	1 = wel	1183	94.3%
7. Etniciteit	0 = Marokkaans	597	47.6%
	1 = Turks	630	50.2%
	2 = Surinaams/Antilliaans	27	2.2%

^aBij continue variabelen is de verdeling uitgedrukt in gemiddelden en standaarddeviaties; ^bBij continue variabelen zijn de minimum en maximum waarden weergegeven in plaats van percentages; ^cn=1083 waarbij de missende waarden niet meegenomen zijn; ^dn=1250 " "; ^en=1092 " "

Tabel 2 geeft een overzicht van de bivariate statistieken weer; in hoeverre er samenhang is tussen alle variabelen die onderdeel zijn van de analyse. Correlaties tussen continue variabelen (leeftijd en opleidingsniveau) en dichotome variabelen (sekse, generatie, zelfde etniciteit resp. en partner) zijn berekend aan de hand van een t-toets die gemiddelde verschillen tussen de groepen toetst. Correlaties tussen twee dichotome variabelen zijn berekend aan de hand van Pearson's correlatie als indicatie voor de phi-coëfficiënt (ϕ). Correlaties tussen continue en categorische variabelen zijn berekend op basis van de ANOVA en de F-toets via de formule $R = \sqrt{1 - \frac{n-k}{n-k+(k-1)F}}$. Tot slot zijn correlaties tussen twee categorische variabelen berekend op basis van Cramer's V. Een uitgebreide toelichting en berekening van de correlaties en verschil in gemiddelden tussen de groepen is te vinden in bijlage 2.

In tabel 2 is te zien dat er een positief verband lijkt te zijn tussen het hebben van co-etnische sociale netwerken in de wijk en het hebben van een partner met dezelfde etniciteit ($r=.23$; $p<0.001$); intensiever contact met co-etnische buurtgenoten lijkt de kans te

vergroten dat iemand een co-etnische partner heeft. Daarnaast lijkt er een significante associatie te zijn tussen zowel *etniciteit* en *co-etnische sociale netwerken in de wijk* ($r=.11$; $p<0.001$) als tussen *etniciteit* en *zelfde etniciteit respondent en partner* ($r=.34$; $p<0.001$). Dit duidt aan dat er verschillen kunnen zijn wat betreft deze twee variabelen tussen de verschillende etnische groepen die onderdeel zijn van dit onderzoek. Daarnaast is aan het positieve verband tussen *generatie migranten* en *co-etnische sociale netwerken in de wijk* te zien dat 1^e generatie migranten doorgaans intensiever contact hebben met co-etnische buurtgenoten ($r=.11$; $p=0.002$) en dat 1^e generatie migranten meer kans hebben om zich in een co-etnisch huwelijk te bevinden ($r=.12$; $p<0.001$). Tot slot is opvallend dat *opleidingsniveau* significant samenhangt met zowel *zelfde etniciteit respondent en partner* (verschil gemiddelde=1.35; $t=3.66$; $p<0.001$) als met *co-etnische sociale netwerken in de wijk* ($r=.16$; $p<0.05$). Respondenten die zich in een co-etnisch huwelijk bevinden zijn gemiddeld iets hoger opgeleid dan respondenten die zich in een interetnisch huwelijk bevinden. Aan de positieve associatie van *co-etnische sociale netwerken in de wijk* en *opleidingsniveau* is te zien dat respondenten die meer contact hebben met co-etnische buurtgenoten gemiddeld genomen hoger opgeleid zijn.

Tabel 2: Samenhangende maten van de in de analyse opgenomen variabelen

	1	2	3	4	5	6
1. Leeftijd	-					
2. Sekse (0=man; 1=vrouw)	2.44 ^{a**}	-				
3. Opleidingsniveau	-.13 ^{**}	.02 ^a	-			
4. Generatie migranten (0=2 ^e generatie; 1=1 ^e generatie)	-5.53 ^{a**}	-.07 ^{b*}	1.24 ^{a**}	-		
5. Co-etnische SN in de wijk (0=vrijwel nooit; 1=geregeld; 2=wekelijks)	.03 ^c	.10 ^{d**}	.16 ^{c*}	.11 ^{d**}	-	
6. Zelfde etniciteit respondent en partner (0=niet; 1=wel)	.688 ^a	-.001 ^b	1.35 ^{a**}	.12 ^{b**}	.23 ^{d**}	-
7. Etniciteit (0=Marokkaans; 1=Turks; 2=Surinaams/Antilliaans)	.08 ^{c*}	.04 ^d	.08 ^{c*}	.10 ^{d**}	.11 ^{d**}	.34 ^{d**}

* significant bij $p<0.05$; ** significant bij $p<0.01$; N ligt tussen de 957 en 1254.

^ahet verschil in gemiddelde leeftijd en opleidingsniveau per groep; ^bPearson's correlatiecoëfficiënt als indicatie voor ϕ ; ^ccorrelatie op basis van ANOVA; ^dCramer's V

3.2 Modevaluatie

Middels het schatten van verschillende modellen is getracht duidelijk te krijgen welke variabelen invloed hebben op iemands keuze voor een partner met dezelfde etniciteit (co-ethnisch huwelijk). Alle modellen zijn weergegeven in tabel 3. Onder in de tabel zijn verschillende waarden te zien die toetsen hoe goed de modellen bij de data passen. De *deviance* waarden, die berekend zijn via de likelihood-ratio toets, geven bijvoorbeeld aan hoeveel schattingsfouten gemaakt zijn; hogere waarden duiden op meer gemaakte fouten. Naarmate er meer variabelen aan de modellen toegevoegd worden neemt de *deviance* waarden af. Model 5 lijkt volgens de *deviance* dan ook het beste bij de data te passen.

De X^2 waarden toetsen of een model significant beter is dan het voorgaande model. Te zien is dat model 1 significant beter is dan het lege model ($X^2(3) = 18.958$; $p < 0.001$), die alleen de afhankelijke variabele bevat. Model 2 is weer significant beter dan model 1 ($X^2(2) = 32.807$; $p < 0.001$) en model 3 op zijn beurt weer beter dan model 2 ($X^2(3) = 65.614$; $p < 0.001$). Echter, model 4 en 5 zijn beiden niet significant beter dan hun voorgaande modellen (respectievelijk $X^2(2) = 0.576$; $p = 0.750$ en $X^2(4) = 5.327$; $p = 0.255$). Het toevoegen van de interactietermen aan het model lijkt niet tot een beter passend model te leiden. Wanneer verondersteld wordt dat het wenselijk is om een klein model te hebben zonder variabelen die weinig toevoegen, kan gezegd worden dat de data het beste bij modellen 1, 2 en 3 passen.

Tot slot laat de Hosmer-Lemeshow toets zien hoe goed elk model bij de data past, door te kijken naar de mate waarin verwachte en geobserveerde waarden overeenkomen. Een hoge p -waarde zou betekenen dat er geen significante verschillen zijn tussen de geobserveerde en verwachte waarden. De toets van elk model is niet-significant, wat betekent dat er geen significante verschillen zijn tussen de geobserveerde en verwachte waarden binnen elk model.

Op basis van de hierboven genoemde bevindingen kan geconcludeerd worden dat model 3 het best past bij de data en dus de best 'fit' heeft. Hoewel modellen 4 en 5 minder schattingsfouten maken en de Hosmer-Lemeshow toets van beiden niet-significant zijn, is model 4 niet significant beter dan model 3 en model 5 niet beter dan model 4. Daarom zal hypothese 1 getoetst worden aan de hand van model 3. Desondanks zullen modellen 4 en 5 wel gebruikt moeten worden om hypothese 2 en 3 te toetsen, aangezien de interactietermen geen onderdeel zijn van model 3.

3.2.1. Assumptie, multicollineariteit & outliers

Om te achterhalen of de resultaten betrouwbaar en generaliseerbaar zijn, is het belangrijk te controleren in hoeverre de data voldoet aan de assumptie die hoort bij een logistische regressieanalyse; de assumptie dat observaties onafhankelijk zijn.

Omdat de dataset verkregen is door het werk van andere onderzoekers, is het belangrijk te kijken hoe zij respondenten geselecteerd hebben. Door hiernaar te kijken kan vastgesteld worden in hoeverre de observaties onafhankelijk zijn. Omdat de NeLLS studie (Tolsma et al., 2014) een over-sample van Turkse en Marokkaanse Nederlanders bevat en mensen uit deze groepen die wonen in kleine dorpen op het platteland niet meegenomen zijn in de studie, zijn de respondenten niet geheel aselekt getrokken maar op een quasi-willekeurige wijze. Om toch onafhankelijke observaties te kunnen verkrijgen zijn 35 gemeenten op quasi-willekeurige wijze geselecteerd door de gemeenten onder te verdelen per regio en urbanisatiegraad. In de uiteindelijke steekproef moesten de vier grote westelijke steden Amsterdam, Rotterdam, Utrecht en Den Haag wel meegenomen worden om een representatieve steekproef van Turkse en Marokkaanse Nederlanders te verkrijgen. Vervolgens zijn respondenten op quasi-willekeurige wijze uit het geboorteregister getrokken op basis van etniciteit en leeftijd, weer met een over-sample van Turkse en Marokkaanse Nederlanders. Op deze manier is getracht onafhankelijke observaties te kunnen verkrijgen.

Daarnaast is het belangrijk dat er geen sprake is van multicollineariteit, dat wil zeggen dat er geen onderlinge samenhang is tussen de variabelen in een model. Wanneer generatie en leeftijd bijvoorbeeld sterk correleren, is het lastig te achterhalen hoeveel invloed zij afzonderlijk van elkaar hebben op het al dan niet hebben van een co-etnisch huwelijk, omdat zij (gedeeltelijk) overlappen. Om te controleren in hoeverre er sprake is van multicollineariteit kan er gekeken worden naar de Variance Inflation Factor (VIF) waarden van alle variabelen in het model. In bijlage 3 tabel 2 is een overzicht gegeven van de VIF-waarden behorend bij de variabelen uit modellen 3, 4 en 5 met een bijbehorende toelichting. Wanneer een grenswaarde van 10 aangenomen wordt, zijn er enkel in model 5 een aantal variabelen die een waarde hebben die hier (ruim) boven vallen. Echter, de variabelen die samenhangen zijn de twee dummy's behorend bij *eticiteit* en de interactietermen. Omdat de twee dummy's samen één variabele vormen en de variabelen die onderdeel zijn van de interactietermen ook één variabele vormen, is het logisch dat deze

VIF-waarden hoog zijn. Toch zullen deze variabelen meegenomen moeten worden in de analyse om de hypothesen te kunnen toetsen. Gevolgen hiervan zijn dat het schatten van de parameters bemoeilijkt zal worden en de betrouwbaarheid van de resultaten in het geding kunnen komen. Hier zal rekening mee gehouden moeten worden bij het trekken van conclusies.

Tot slot is het belangrijk om te achterhalen in hoeverre de data invloedrijke punten bevat. De Leverage waarden behorend bij alle cases kunnen inzicht geven in welke punten mogelijk invloedrijk zijn. Deze waarden laten namelijk zien in hoeverre de cases invloed uitoefenen op de fit van de modellen. Wanneer één case veel invloed zou hebben op de modelfit, zou deze case bepalend kunnen zijn voor de andere cases. De bovengrens van de Leverage waarden kan berekend worden via de formule $3 \cdot p/n$. Hierbij staat p voor het aantal variabelen in het model en n voor de steekproefgrootte. Waarden boven deze grens worden gezien als invloedrijke punten en zijn dus mogelijk outliers. Voor model 5 is dit $3 \cdot 14/953 = 0.044$. Figuur 1 uit bijlage 3 laat zien dat er een aantal cases zijn die boven de grens van 0.044 vallen, namelijk 69 cases. Deze cases zouden invloed kunnen hebben op de geschatte parameters en de fit van de modellen beïnvloed kunnen hebben. Alle cases zullen echter wel meegenomen worden in de analyse. Leverage waarden zijn namelijk vooral relevante waarden om outliers te detecteren van continue variabelen en minder betekenisvol bij binaire logistische regressies. Dichotome variabelen zoals de afhankelijke in dit onderzoek kennen immers maar twee scores, namelijk 0 of 1. Daar komt bij dat de outliers in binaire logistische regressies juist betekenisvol kunnen zijn, doordat zij, zoals in dit onderzoek het geval is, anders scoren dan de meerderheid op bepaalde variabelen en hierdoor scheve verdelingen kunnen voorkomen (zie bijlage 3 figuur 2 voor een verdeling en verdere toelichting).

3.3 Bespreking belangrijke resultaten en hypothesetoetsing

Zoals net gesteld kunnen de hypothesen getoetst worden aan de hand van de modellen en bevat elk model weer andere resultaten. Zo laat model 1 uit tabel 3 zien wat de invloed van de drie controlevariabelen zijn op de afhankelijke variabele. Aan de negatieve helling van opleidingsniveau ($^{model1}b = -0.213$; $p < 0.001$) is te zien dat de kans op een co-etnisch huwelijk kleiner is voor respondenten die een hoger opleidingsniveau hebben. Deze variabele blijft in elk model een negatieve helling behouden en blijft significant ($p < 0.05$).

Daarnaast is aan de odds-ratio (OR=2.325) van *generatie* ($b=0.844$; $p=0.020$) in model 3 te zien dat 1^e generatie migranten meer kans hebben op een co-etnisch huwelijk dan 2^e generatie migranten. Wanneer de odds-ratio waarde kleiner dan 1 zou zijn, zou de kans dat 1^e generatie migranten co-etnisch getrouwd zijn kleiner zijn dan de kans van 2^e generatie migranten. Wanneer de odds-ratio 1 zou zijn, zou er geen verschil zijn tussen beide groepen. De parameters van *generatie* blijven constant in de daaropvolgende twee modellen; de hellingen blijven positief (^{model4} $b=1.124$ en $p=0.058$; ^{model5} $b=1.185$ en $p=0.050$) en de odds-ratio waarden boven 1 (^{model4}2.712; ^{model5}2.822). Echter, de resultaten zijn (net) niet langer meer significant wanneer de interactietermen aan het model worden toegevoegd. Dit zou kunnen betekenen dat er alleen significante verschillen zijn tussen de 1^e en 2^e generatie afhankelijk van de mate waarin zij contact hebben met co-etnische buurtgenoten (hypothese 2). Verderop in deze sectie zal duidelijk worden of dit het geval is.

Zowel de helling behorend bij *etniciteit Marokkaans* als bij *etniciteit Turks* blijft daarentegen wel in alle modellen significant ($p \leq 0.005$) en de odds-ratio waarden boven 1. Aan de odds-ratio waarden van bijvoorbeeld *etniciteit Turks* in model 3 is te zien dat de odds op een co-etnisch huwelijk 59.381 ($b=4.084$; $p < 0.001$) keer groter is voor Turkse Nederlanders dan voor Surinaamse/Antilliaanse Nederlanders. De odds is 11.412 ($b=2.435$; $p < 0.001$) keer groter voor Marokkaanse Nederlanders dan voor Surinaamse/Antilliaanse Nederlanders, wat laat zien dat Surinaamse/Antilliaanse Nederlanders een kleinere kans hebben om zich in een co-etnisch huwelijk te bevinden dan de andere twee groepen.

Tevens is te zien dat de kans op een co-etnisch huwelijk groter is voor mensen die geregeld of wekelijks contact hebben met co-etnische buurtgenoten dan voor mensen die vrijwel nooit ($\leq 1x$ per jaar) contact hebben met co-etnische buurtgenoten (referentiegroep); de hellingen uit model 3 zijn positief en significant (respectievelijk $p=0.004$ en $p < 0.001$) en de odds-ratio waarden > 1 (respectievelijk 3.532 en 5.024). Deze bevindingen zijn in lijn met hypothese 1, die stelt dat mensen sneller co-etnisch trouwen wanneer er een hogere mate van co-etnische sociale netwerken in hun wijk is.

De parameters behorend bij de interactietermen van *generatie* en *co-etnische sociale netwerken in de wijk* in model 4 lijken daarentegen niet in lijn te zijn met hypothese 2. De helling van één interactietermen is negatief ($b=-0.165$; $se=0.912$; $p=0.858$) en de odds-ratio waarden van beide termen zijn < 1 (^{interactieterm1}OR=0.848 en ^{interactieterm2}OR=0.553), wat laat zien dat de invloed van co-etnische sociale netwerken in de wijk op iemands keuze voor

een co-etnisch huwelijk sterker is voor 2^e generatie migranten dan voor 1^e generatie migranten. Dit is het tegenovergestelde van wat hypothese 2 stelt, namelijk dat dit verband sterker is voor 1^e generatie migranten. De hellingen zijn echter relatief klein, niet-significant en de standaardfouten relatief groot. Het model lijkt de helling van de interactievariabelen moeilijk te kunnen schatten. Wat wel duidelijk is, is dat er significante verschillen zijn tussen respondenten die vrijwel nooit en respondenten die wekelijks contact hebben met co-etnische buurtgenoten, gecontroleerd voor de interactietermen ($^{model4}b=2.026; p=0.003$). Deze verschillen blijven, ook als er gecontroleerd wordt voor de interactietermen. Dit laat nogmaals zien dat er geen significante verschillen gevonden zijn tussen beide generaties wat betreft het hoofdeffect, waardoor hypothese 2 niet aangenomen kan worden.

Daarnaast zijn aan drie negatieve hellingen van de interactietermen *etniciteit: Turks/Marokkaans* en *co-etnische sociale netwerken in de wijk* te zien dat het hoofdeffect minder sterk is voor Turkse en Marokkaanse Nederlanders dan voor Surinaamse/Antilliaanse Nederlanders (variërend van $b=-1.488$ tot $b=-0.699$). Dit is het tegenovergestelde van wat hypothese 3 stelt, namelijk dat dit effect sterker zou zijn voor Turkse en Marokkaanse Nederlanders dan voor Surinaamse/Antilliaanse Nederlanders. Echter, de hellingen zijn klein, niet significant en de standaardfouten groot (bijvoorbeeld $interactieterm1 \times Marokkaans b=-0.699; se=1.488; p=0.629$). Eén helling is positief, namelijk de helling van Turkse Nederlanders die geregeld contact hebben met co-etnische buurtgenoten ($b=16.980; p=0.997$). Maar, de standaardfout is erg groot ($se=4173.878$) en de helling niet significant. Over deze parameter kunnen dan ook geen betrouwbare uitspraken gedaan worden. Deze bevindingen, en het gegeven dat de significante verschillen tussen respondenten die vrijwel nooit en zij die wekelijks contact hebben met co-etnische buurtgenoten blijvend zijn ook wanneer er gecontroleerd wordt voor de interactietermen, laten zien dat de resultaten hypothese 3 niet ondersteunen en deze hypothese dus niet aangenomen kan worden. Hoewel Turkse en Marokkaanse Nederlanders volgens de resultaten vaker co-etnisch trouwen dan Surinaamse/Antilliaanse Nederlanders, is het nog onduidelijk in hoeverre contact met co-etnische buurtgenoten hierbij een rol speelt.

Tabel 3: Resultaten van de hiërarchische logistische regressieanalyse met als afhankelijke variabele keuze voor partner zelfde etniciteit (n=953)

	Model 1			Model 2			Model 3			Model 4			Model 5		
	b(SE)	Odds-ratio	p	b(SE)	Odds-ratio	p	b(SE)	Odds-ratio	p	b(SE)	Odds-ratio	p	b(SE)	Odds-ratio	p
Constante	5.279 (1.005)	196.262	<0.001	3.716 (1.078)	41.091	0.001	0.021 (1.325)	1.022	0.987	-0.366 (1.432)	0.693	0.798	-1.048 (1.770)	0.351	0.554
Leeftijd	-0.036 (0.024)	0.965	0.138	-0.035 (0.025)	0.965	0.157	-0.030 (0.029)	0.970	0.300	-0.027 (0.029)	0.983	0.357	-0.028 (0.029)	0.973	0.347
Sekse (1=vrouw)	-0.206 (0.285)	0.814	0.469	-0.258 (0.296)	0.776	0.369	-0.059 (0.329)	0.943	0.858	-0.018 (0.334)	0.983	0.958	0.010 (0.336)	1.010	0.977
Opleidingsniveau	-0.213 (0.053)	0.808	<0.001	-0.170 (0.054)	0.842	0.001	-0.141 (0.060)	0.869	0.019	-0.137 (0.060)	0.872	0.023	-0.145 (0.062)	0.865	0.019
(1) Co-etnische sociale netwerken in de wijk; geregeld maar <1x per week contact				1.226 (0.386)	3.408	0.001	1.262 (0.437)	3.532	0.004	1.373 (0.724)	3.948	0.058	-1.692 (1.572)	5.429	0.282
(2) Minstens 1x per week contact				1.941 (0.327)	6.963	<0.001	1.614 (0.362)	5.024	<0.001	2.026 (0.671)	7.587	0.003	3.352 (1.454)	28.560	0.021
Generatie (1=1 ^e generatie)							0.844 (0.364)	2.325	0.020	1.124 (0.594)	3.077	0.058	1.185 (0.604)	3.269	0.050
Etniciteit: Marokkaans (Surinaams/Antillia ans=ref)							2.435 (0.490)	11.412	<0.001	2.507 (0.503)	12.270	<0.001	3.340 (1.182)	28.215	0.005
Etniciteit: Turks (Surinaams/Antillia ans=ref)							4.084 (0.574)	59.381	<0.001	4.136 (0.583)	62.548	<0.001	4.642 (1.288)	103.711	<0.001
Co-etnische sn in de wijk (1) X generatie										-0.165 (0.912)	0.848	0.857	-0.043 (0.974)	0.958	0.965

Co-etnische sn in de wijk (2) X generatie						0.593 (0.809)	0.553	0.464	-0.589 (0.805)	0.555	0.464
Co-etnische sn in de wijk (1) X Marokkaans									-0.699 (1.448)	0.497	0.629
Co-etnische sn in de wijk (2) X Marokkaans									-1.488 (1.423)	0.226	0.296
Co-etnische sn in de wijk (1) X Turks									16.980 (4173.878) ^b	>23milj	0.997
Co-etnische sn in de wijk (2) X Turks									-1.392 (1.554)	0.248	0.370
<i>Deviance</i>	412.651		379.844		314.230		313.654			308.327	
<i>X²-toets</i>	18.958 ^a	<0.001	32.807	<0.001	65.614	<0.001	0.576	0.750	5.327		0.255
Hosmer-Lemeshow toets	10.206	0.251	7.960	0.437	3.443	0.904	5.923	0.656	2.090		0.978
<i>n</i>	953		953		953		953		953		

^aTen opzichte van het lege model. ^bDeze (interactie)variabele bevat onvoldoende betrouwbare informatie om uitspraken over te kunnen doen. De geschatte parameters behorende bij deze variabele zullen verder dan ook niet uitgebreid besproken worden.

H4: Conclusie en discussie

Conclusie

Dit onderzoek had als doel om inzichtelijk te maken wat van invloed kan zijn op iemands keuze voor een co-etnisch huwelijk. De kans dat mensen met verschillende etnische achtergronden met elkaar trouwen is namelijk groter dan decennia geleden, omdat er in Nederland een multiculturele samenleving is ontstaan door de jarenlange toestroom van migranten. Het aangaan van een dergelijk huwelijk kan worden gezien als de laatste stap die een persoon met een migratieachtergrond kan zetten om volledig te integreren (Tolsma et al., 2007) en zou ervoor kunnen zorgen dat grenzen tussen etnische groepen vervagen en vooroordelen over andere etnische groepen verminderen. Echter, slechts 16% van alle voltrokken huwelijken in 2016 was tussen iemand met een migratieachtergrond en een Nederlander. Vooral Turkse en Marokkaanse Nederlanders kiezen er vaker voor om te trouwen met iemand die dezelfde etniciteit heeft vergeleken met Surinaamse en Antilliaanse Nederlanders; 80% van zowel de Turkse als de Marokkaanse Nederlanders heeft een partner met dezelfde etniciteit (CBS, 2020). Voor Surinaamse Nederlanders is dit slechts 50% en voor Antilliaanse Nederlanders zelfs maar 25% (CBS, 2020). Daarnaast zouden voornamelijk 1^e generatie migranten sneller een co-etnisch huwelijk aangaan dan een interetnisch huwelijk, omdat zij ten opzichte van de 2^e generatie meer gesocialiseerd zijn in hun geboorteland en de Nederlandse taal veelal minder goed spreken (SCP, 2020).

Verschillende factoren kunnen van invloed zijn op iemands keuze voor een co-etnisch huwelijk, zo ook de factor die centraal stond in deze studie; de aanwezigheid van co-etnische sociale netwerken in de wijk. Zowel het nabijheidseffect als het principe van ruimtelijke homogamie laten namelijk zien dat mensen sneller bevriend raken met iemand die dicht bij hen woont en sneller geneigd zijn te trouwen met mensen die dichtbij wonen (Festinger, 1973; Haandrikman, 2018). De onderzoeksvraag luidde dan ook: *“Hangt een hoge mate aan co-etnische sociale netwerken in de wijk samen met individuele partnerkeuze en verschilt dit voor verschillende etnische groepen en voor 1^e en 2^e generatie migranten?”* Om deze onderzoeksvraag te kunnen beantwoorden zijn 3 hypothesen opgesteld die getoetst zijn aan de hand van een logistische regressieanalyse.

Hypothese 1 stelde dat een hogere mate van co-etnische sociale netwerken in de wijk gepaard zou gaan met een grotere kans dat iemand uit die wijk kiest voor een partner die dezelfde etniciteit heeft. Uit de resultaten is gebleken dat mensen die wekelijks of geregeld contact hebben met co-etnische buurtgenoten, meer kans hebben om zich in een co-etnisch huwelijk te bevinden dan mensen die vrijwel nooit contact hebben met co-etnische buurtgenoten. Dit is in lijn met de theoriesectie, die stelt dat wanneer er meer ontmoetingsmogelijkheden zijn tussen mensen de kans op een huwelijk ook groter wordt (Kalmijn, 1998).

Hypothese 2 stelde dat het positieve effect van de mate van co-etnische sociale netwerken in de wijk op de kans dat iemand uit die wijk kiest voor een partner die dezelfde etniciteit heeft, sterker is voor 1^e generatie migranten dan voor 2^e generatie migranten. De 2^e generatie zou namelijk meer ontmoetingsmogelijkheden hebben met bijvoorbeeld Nederlanders buiten de wijk om, doordat zij vaak naar Nederlandse scholen gaan en de taal beter spreken dan de 1^e generatie. Maar, uit de resultaten is gebleken dat er geen duidelijke verschillen zijn gevonden tussen de 1^e en 2^e generatie wat betreft de invloed van co-etnische sociale netwerken in de wijk op iemands partnerkeuze. Dit zou kunnen komen doordat de groep 2^e generatie migranten klein is in vergelijking met de groep 1^e generatie migranten, wat op zijn beurt de betrouwbaarheid van de resultaten beïnvloed kan hebben. Dit zal in de discussiesectie uitgebreider besproken worden. Wel laten de resultaten zien dat 1^e generatie migranten meer kans hebben om zich in een co-etnisch huwelijk te bevinden dan de 2^e generatie.

Tot slot stelde hypothese 3 dat het positieve effect van de mate van co-etnische sociale netwerken in de wijk op de kans dat iemand uit die wijk kiest voor een partner die dezelfde etniciteit heeft, sterker is voor Turkse en Marokkaanse Nederlanders dan voor Surinaamse en Antilliaanse Nederlanders. Surinaamse en Antilliaanse Nederlanders spreken namelijk doorgaans beter Nederlands en voelen zich vaker thuis in Nederland dan Turkse en Marokkaanse Nederlanders (CBS, 2022b). Hierdoor spelen co-etnische sociale netwerken in de wijk wellicht een minder significante rol in hun partnerkeuze dan voor Turkse en Marokkaanse Nederlanders. Uit de resultaten is echter gebleken dat er geen duidelijke verschillen zijn gevonden tussen de etnische groepen wat betreft de invloed van co-etnische sociale netwerken in de wijk op iemands partnerkeuze. Dit zou kunnen komen doordat er een sterke samenhang is tussen de gebruikte variabelen (multicollineariteit), waardoor

uitspraken die gebaseerd zijn op deze resultaten minder betrouwbaar zijn. Daarnaast zou het ook kunnen komen doordat de groep Surinaamse/Antilliaanse Nederlanders erg klein is in de steekproef en de kans op significante resultaten hierdoor kleiner is. Desondanks kan uit de resultaten wel opgemaakt worden dat Turkse en Marokkaanse Nederlanders meer kans hebben om zich in een co-etnisch huwelijk te bevinden dan Surinaamse en Antilliaanse Nederlanders.

Discussie

Dit onderzoek heeft echter enkele beperkingen waardoor sommige resultaten en conclusies niet geheel betrouwbaar zijn.

Zoals net al naar voren kwam is er sprake van multicollineariteit, wat de toetsing van hypothese 3 bemoeilijkt heeft. Hierdoor zijn de resultaten minder betrouwbaar en is het onduidelijk of de invloed van co-etnische sociale netwerken in de wijk op iemands partnerkeuze sterker is voor Turkse en Marokkaanse Nederlanders of dat er alleen een verband is tussen etniciteit en iemands partnerkeuze.

Verder zijn een aantal groepen significant groter dan andere groepen. Zo is de groep 1^e generatie migranten veel groter dan de groep 2^e generatie migranten (respectievelijk 82.6% en 17.4%) en zijn vrijwel alle respondenten getrouwd met iemand die dezelfde etniciteit heeft (94.3%). Dit zou de betrouwbaarheid van de resultaten kunnen aantasten; wanneer een groep klein is kan één respondent al veel invloed hebben op de resultaten. Zo is het aandeel Surinaamse/Antilliaanse Nederlanders in de dataset erg klein ($n=27$, 2.2%), wat kan komen doordat dit onderzoek gebaseerd is op secundaire gegevens die een oversample van Turkse en Marokkaanse Nederlanders bevat. Desondanks is het aandeel Surinaamse en Antilliaanse Nederlanders uit de dataset die interetnisch getrouwd zijn (55.6%) ongeveer even groot als het aandeel uit de gehele bevolking (respectievelijk 50% en 75%). Maar, om nauwkeurige resultaten te verkrijgen is het van belang dat in vervolgonderzoek de te onderzoeken etnische groepen een afspiegeling zijn, in aantallen, van het aandeel uit de gehele bevolking. Zo is het aandeel mensen met een Surinaamse en Antilliaanse migratieachtergrond van alle mensen in Nederland met een migratieachtergrond ruim 12% (CBS Statline, 2022), terwijl dit aandeel slechts 2.2% is in dit onderzoek.

Daar komt bij dat het moeilijk kan zijn om etniciteit te meten, omdat het verschillend gedefinieerd kan worden en daardoor ook verschillend gemeten kan worden. Zo is de etniciteit van de respondent op een net iets andere manier gemeten dan de etniciteit van diens partner (zie bijlage 1). Het hanteren van eenzelfde definitie zou kunnen leiden tot een nauwkeuriger onderzoek en is dan ook van belang voor vervolgonderzoek.

Daarnaast kan de invloed van de wijk wellicht beter onderzocht worden aan de hand van een steekproef bestaande uit verloofde en pasgetrouwde mensen met een migratieachtergrond. Omdat mensen die al lang getrouwd zijn wellicht pas in hun wijk zijn gaan wonen toen zij al getrouwd waren, heeft de wijk geen invloed gehad op hun partnerkeuze. Ook kan het waardevol zijn om aan respondenten te vragen waar zij hun partner ontmoet hebben, om zo ook beter de invloed van de wijk te kunnen achterhalen. Onderscheid maken tussen jongere en oudere mensen met een migratieachtergrond kan hierin interessant zijn, aangezien uit de literatuur is gebleken dat ouderen een grotere kans zouden hebben om zich in een co-etnisch huwelijk te bevinden dan jongeren (Kalmijn & van Tubergen, 2007). Om dit te kunnen toetsen en een goede vergelijking te kunnen maken tussen de groepen is het nodig om ook 50+ers mee te nemen in onderzoek.

Een laatste beperking van dit onderzoek is dat er alternatieve verklaringen kunnen zijn voor waarom sommige mensen wel een co-etnisch huwelijk aangaan en anderen niet. Zo toont de literatuur aan dat religie een grote invloed kan hebben op iemands keuze voor een co-etnisch huwelijk; de sociale afstand tussen mensen die aan verschillende religies verbonden zijn zou groot zijn en bepaalde belangrijke waarden en normen worden wellicht niet gedeeld (Kalmijn & van Tubergen, 2010). Het welbekende gezegde "twee geloven op een kussen, daar slaapt de duivel tussen" stelt dat een interreligieus huwelijk tot moeilijkheden kan leiden. Maar, omdat vrijwel alle respondenten islamitisch zijn en een enkeling een ander geloof aanhangt was het niet mogelijk de groepen met elkaar te kunnen vergelijken.

Ondanks deze beperkingen is door dit onderzoek nogmaals duidelijk geworden dat contact met co-etnische buurtgenoten de kans op een co-etnisch huwelijk kan vergroten. Hoewel mensen tegenwoordig met mensen van over de hele wereld in contact staan door apps als Whatsapp en Skype en relaties ontstaan via dating apps, blijft iemands directe leefomgeving van invloed op diens partnerkeuze. Tot welke generatie migranten iemand behoort is daarnaast ook van invloed; de kans op een co-etnisch huwelijk is groter voor 1^e

generatie migranten vergeleken met 2^e generatie migranten. In hoeverre het contact dat zij hebben met co-etnische buurtgenoten daarin een belangrijke rol speelt is alleen nog onduidelijk. Het zou daarom ook relevant zijn om dit nader te onderzoeken. Belangrijk voor vervolgonderzoek is dat de steekproef aansluit bij het doel van het onderzoek en er geen ondervetegenwoordigde groepen zijn. Daarnaast zou het relevant zijn om alternatieve factoren als religie mee te nemen in onderzoek, om zo ook betrouwbaardere uitspraken te kunnen doen en een completer beeld te krijgen.

Tot slot zou het relevant kunnen zijn om beleid op te stellen dat contact tussen Nederlanders en mensen met een migratieachtergrond faciliteert. Zoals al eerder is gesteld kan een interetnisch huwelijk gezien worden als de laatste stap die iemand met een migratieachtergrond kan zetten om volledig te integreren. Het begint echter bij ontmoetingsmogelijkheden; hoe meer ontmoetingsmogelijkheden er zijn tussen twee mensen, des te groter de kans dat zij elkaar als partner zullen kiezen (Kalmijn, 1998). Om de integratie van bijvoorbeeld Turkse en Marokkaanse Nederlanders te bevorderen zou de gemeente meer kunnen investeren in bijvoorbeeld buurthuizen, om het zo een aantrekkelijkere plek te maken om naartoe te gaan. Omdat buurthuizen doorgaans voornamelijk bezocht worden door mensen met een migratieachtergrond (Veld, 2004) zouden buurthuizen diverse soorten cursussen en activiteiten kunnen aanbieden die ook Nederlanders aan kunnen trekken, zoals activiteiten die gerelateerd zijn aan muziek of kunst. Gemeenschappelijk interesses, zoals muziek en kunst, kunnen namelijk een middel zijn om mensen met verschillende etnische achtergronden met elkaar in contact te brengen (Geugies, 2009). Dit zou de integratie van Turkse en Marokkaanse Nederlanders kunnen bevorderen, doordat de kloof tussen hen en Nederlanders op deze manier kleiner kan worden.

Literatuurlijst

- Batool, S., & Malik, N. I. (2010). Role of attitude similarity and proximity in interpersonal attraction among friends. *International Journal of Innovation, Management and Technology*, 1(2), 142-146.
- van den Berg, P., & Timmermans, H. (2015). A multilevel path analysis of social networks and social interaction in the neighbourhood. *Region*, 2(1), 55-66.
- Bolt, G., & van Kempen, R. (2010). Ethnic Segregation and Residential Mobility: Relocations of Minority Ethnic Groups in the Netherlands. *Journal of Ethnic and Migration Studies*, 36(2), 333-354. <https://doi.org/10.1080/13691830903387451>
- Carol, S. (2014). Like Will to Like? Partner Choice among Muslim Migrants and Natives in Western Europe. *Journal of Ethnic and Migration Studies*, 42(2), 261–276. <https://doi.org/10.1080/1369183x.2014.963037>
- CBS. (2017, 20 januari). *Gemengd samenwonen verschilt per migratieachtergrond*. Geraadpleegd op 1 april 2022, van <https://www.cbs.nl/nl-nl/nieuws/2017/03/gemengd-samenwonen-verschilt-per-migratieachtergrond>
- CBS. (2020, 11 november). *Bevolking - Jaarrapport Integratie 2020*. Bevolking - Jaarrapport Integratie 2020 | CBS. Geraadpleegd op 2 maart 2022, van <https://longreads.cbs.nl/integratie-2020/bevolking/>
- CBS. (2022a, 16 februari). *Hoe verschillen arbeid en inkomen naar migratieachtergrond?* Geraadpleegd op 1 april 2022, van <https://www.cbs.nl/nl-nl/dossier/dossier-asiel-migratie-en-integratie/hoe-verschillen-arbeid-en-inkomen-naar-migratieachtergrond->
- CBS. (2022b, 25 februari). *Hoe verschillen opleiding en schoolkeuze naar migratieachtergrond?* Geraadpleegd op 1 april 2022, van <https://www.cbs.nl/nl-nl/dossier/dossier-asiel-migratie-en-integratie/hoe-verschillen-opleiding-en-schoolkeuze-naar-migratieachtergrond->

- CBS Statline. (2011). [Dataset]. Centraal Bureau voor de Statistiek.
<https://opendata.cbs.nl/#/CBS/nl/dataset/80270ned/table>
- CBS Statline. (2022). [Dataset]. Centraal Bureau voor de Statistiek.
<https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70787ned/table?fromstatweb>
- Festinger, L. (1973). *Leon festinger*. BSIP Ferranti Limited.
- Fong, P., Cruwys, T., Haslam, C., & Haslam, S. A. (2019). Neighbourhood identification and mental health: How social identification moderates the relationship between socioeconomic disadvantage and health. *Journal of Environmental Psychology, 61*, 101-114. <https://doi.org/10.1016/j.jenvp.2018.12.006>
- Geugies, C. J. (2009). *Het zijn ook gewoon mensen: een onderzoek naar de werking van interetnisch contact* (Master's thesis, University of Twente). Geraadpleegd van <https://essay.utwente.nl/60214/>
- Haandrikman, K. (2018). Partner choice in Sweden: How distance still matters. *Environment and Planning A: Economy and Space, 51*(2), 440–460.
<https://doi.org/10.1177/0308518x18786726>
- Hartung, A., Vandezande, V., Phalet, K., & Swyngedouw, M. (2011). Partnership preferences of the Belgian second generation: Who lives with whom? *Advances in Life Course Research, 16*(4), 152–163. <https://doi.org/10.1016/j.alcr.2011.09.001>
- Kalmijn, M. (1998). Inter-marriage and homogamy: Causes, patterns, trends. *Annual Review of Sociology, 24*(1), 395–421. <https://doi.org/10.1146/annurev.soc.24.1.395>
- Kalmijn, M., & Van Tubergen, F. (2007). Ethnic intermarriage in the Netherlands: confirmations and refutations of accepted insights. *European Journal of Population, 22*(4), 371–397. <https://doi.org/10.1007/s10680-006-9105-3>
- Kalmijn, M., & van Tubergen, F. (2010). A comparative perspective on intermarriage: Explaining differences among national-origin groups in the United States. *Demography, 47*(2), 459–479. <https://doi.org/10.1353/dem.0.0103>

- Keijzer, K. (2021). *Liking* [Powerpoint-slides]. Faculteit Gedrags- en Maatschappijwetenschappen, Rijksuniversiteit Groningen. Geraadpleegd op 4 maart 2022, van <https://signon.rug.nl/>
- SCP. (2020, 18 december). *Arbeidsmarktpositie tweede generatie afgelopen jaren sterk verbeterd*. Geraadpleegd op 21 mei 2022, van <https://www.scp.nl/actueel/nieuws/2020/12/17/arbeidsmarktpositie-tweede-generatie-afgelopen-jaren-sterk-verbeterd>
- Tolsma, J., Lubbers, M., & Coenders, M. (2007). Ethnic Competition and Opposition to Ethnic Inter-marriage in the Netherlands: a Multi-Level Approach. *European Sociological Review*, 24(2), 215–230. <https://doi.org/10.1093/esr/jcm047>
- Tolsma, J., Kraaykamp, G., Graaf, P.M., Kalmijn, M., Monden, C.W.S. (2014). *The NETHERLANDS Longitudinal Lifecourse Study (NELLS, Panel)*. Radboud University Nijmegen, Tilburg University & University of Amsterdam, Netherlands.
- Veld, T.W.M. (2004). *Interetnische contacten op portieken en in voetbalclubs: De ervaringen van allochtone en autochtone bewoners uit de Schilderswijk en Bouwlust*. Geraadpleegd van <https://hdl.handle.net/1765/33619>
- van Zantvliet, P. I., Kalmijn, M., & Verbakel, E. (2014). Early Partner Choices of Immigrants: The Effect of Preferences, Opportunities and Parents on Dating a Native. *Journal of Ethnic and Migration Studies*, 41(5), 772–794. <https://doi.org/10.1080/1369183x.2014.948391>

Bijlagen

Bijlage 1: Gebruikte variabelen en hun bewerkingen

In deze bijlage zijn de verdelingen van de oorspronkelijke variabelen en bewerkte variabelen die onderdeel zijn van dit onderzoek te zien met de bijbehorende syntax. Sommige oorspronkelijke variabelen zijn op verschillende manieren bewerkt, bijvoorbeeld de zelfgerapporteerde etniciteit van de respondenten, om zo verschillende concepten te kunnen meten.

Echter, om de oorspronkelijke en bewerkte variabelen met elkaar te kunnen vergelijken is het belangrijk dat de gegevens betrekking hebben op dezelfde steekproef. Omdat dit onderzoek specifiek gericht is op gehuwde Turkse, Marokkaanse of Surinaamse/Antilliaanse Nederlanders zijn alle respondenten die niet tot één van deze groepen behoort uit de dataset gehaald.

Syntax:

FILTER OFF.	FILTER OFF.
USE ALL.	USE ALL.
SELECT IF (w1cethnic <= 6).	SELECT IF (w1fb10 = 1).
EXECUTE.	EXECUTE.

Omdat de gebruikte *eticiteit* (w1cethnic) variabele in plaats van een aparte categorie voor Surinaamse/Antilliaanse respondenten alleen de categorieën "1^e generatie niet-westersen" en "2^e generatie niet-westersen" heeft, is het nodig deze categorieën verder te definiëren. Hiervoor is gekeken naar het geboorteland van de respondent en diens ouders. Wanneer de respondent geboren is in bijvoorbeeld Suriname is diegene beschouwd als een 1^e generatie Surinaamse, maar wanneer één of beide ouders is geboren in Suriname is de respondent beschouwd als een 2^e generatie Surinaamse. Aruba (score 16) is beschouwd als onderdeel van de Nederlandse Antillen (score 5).

Syntax:

```
COMPUTE Antilliaans_1egen=(w1fa4a = 5 or w1fa4a = 16).  
EXECUTE.
```

```
COMPUTE Antilliaans_2egen=(w1fa4b = 5 or w1fa4b = 16) OR (w1fa4c = 5 or w1fa4c = 16).  
EXECUTE.
```

```
COMPUTE Surinaams_1egen=w1fa4a = 4.  
EXECUTE.
```

```
COMPUTE Surinaams_2egen=(w1fa4b = 4 or w1fa4c = 4).  
EXECUTE.
```

De NeLLS studie (Tolsma et al., 2014) hanteert de volgende definitie van etniciteit: "De exacte herkomst wordt bepaald door het geboorteland van de moeder als beide ouders (of alleen de moeder) buiten Nederland zijn geboren. Als de moeder in Nederland is geboren, wordt de herkomst bepaald door het geboorteland van de vader. Een persoon is van Nederlandse afkomst als beide ouders in Nederland zijn geboren, ongeacht het eigen geboorteland." Desondanks is ervoor gekozen om een respondent als Surinaams/Antilliaans te bestempelen wanneer in ieder geval één van beide ouders in deze landen geboren is, in plaats van het geboorteland van de moeder als lijdend te zien. Werd dit niet gedaan, dan was de groep nog kleiner geworden van die nu al is ($n=27$) en kon deze groep niet meegenomen worden in de analyse.

Vervolgens zijn de andere niet-westerse respondenten uit de dataset gehaald, middels de volgende syntax:

```
FILTER OFF.  
USE ALL.  
SELECT IF ((w1cethnic <= 4) OR (Surinaams_1egen = 1) OR (Surinaams_2egen = 1) OR  
(Antilliaans_1egen = 1) OR (Antilliaans_2egen = 1)).  
EXECUTE.
```

Co-etnische sociale netwerken in de wijk (onafhankelijke variabele)

Deze variabele is ontstaan door het bij elkaar voegen van verschillende items, namelijk de zelfgerapporteerde etniciteit van de respondenten, het geboorteland van de respondenten en het geboorteland van diens ouders en tot slot de items die aangeven in welke mate de respondenten contact hebben met buurtgenoten van Turkse, Marokkaanse of Surinaamse/Antilliaanse afkomst.

Er is gebruik gemaakt van de zelfgerapporteerde etniciteit van de respondenten, om zo te achterhalen of het contact met buurtgenoten co-etnisch is of niet. De verdeling van dit item is weergegeven in tabel 1.

Syntax:

FREQUENCIES VARIABLES=1cethnic
/ORDER=ANALYSIS.

Tabel 1: Frequentieverdeling zelfgerapporteerde etniciteit van de respondenten

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Moroccan, 1st gen	513	39,4	39,4	39,4
	2 Moroccan, 2nd gen	85	6,5	6,5	45,9
	3 Turkish, 1st gen	507	38,9	38,9	84,8
	4 Turkish, 2nd gen	128	9,8	9,8	94,6
	5 Non West, 1st gen	60	4,6	4,6	99,2
	6 Non West, 2nd gen	10	,8	,8	100,0
	Total	1303	100,0	100,0	

Te zien is dat het merendeel 1^e generatie Marokkaans of 1^e generatie Turks is. De 2^e generatie groep is aanzienlijk kleiner dan de 1^e generatie. Daarnaast is te zien dat de groep niet-westersen klein is ten opzichte van de andere groepen, wat begrijpelijk is aangezien de NeLLS studie voornamelijk gericht is op Turkse en Marokkaanse Nederlanders.

Om één etniciteit variabele te maken zijn de eerder gemaakte 4 items samengevoegd met de zelfgerapporteerde etniciteit van de respondenten, waarbij de 1^e generatie samengenomen is met de 2^e generatie. Categorieën 5 en 6 (niet-westersen) zijn hierbij buiten beschouwing gelaten. Daarnaast zijn de groepen Surinaamse en Antilliaanse respondenten samengenomen, omdat beide groepen erg klein zijn (respectievelijk $n=24$ en $n=3$). De bijbehorende output en syntax zijn:

Tabel 2: Frequentieverdeling nieuwe etniciteit variabele

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00 Marokkaans	597	47,6	47,6	47,6
	1,00 Turks	630	50,2	50,2	97,8
	2,00 Antilliaans/Surinaams	27	2,2	2,2	100,0
	Total	1254	100,0	100,0	

```
IF (w1cethnic = 1) Ethnic_new = 0.
IF (w1cethnic = 2) Ethnic_new = 0.
IF (w1cethnic = 3) Ethnic_new = 1.
IF (w1cethnic = 4) Ethnic_new = 1.
IF (Antilliaans_1egen = 1) Ethnic_new = 2.
IF (Antilliaans_2egen = 1) Ethnic_new = 2.
IF (Surinaams_1egen = 1) Ethnic_new = 3.
IF (Surinaams_2egen = 1) Ethnic_new = 3.
EXECUTE.
```

```
RECODE Ethnic_new (0=0) (1=1) (2=2) (3=2) INTO Ethnic_final.
VARIABLE LABELS Ethnic_final 'Drie groepen; Marokkaans, Turks, Antilliaans+Surinaams'.
EXECUTE.
```

Vervolgens is gekeken naar hoe intensief de respondenten contact hebben met co-etnische buurtgenoten (variërend van 1 = "(bijna) elke dag" tot 7 = "nooit"). Hiervoor zijn drie items vergeleken met de etniciteit van de respondenten. Geen enkele respondent heeft voor de

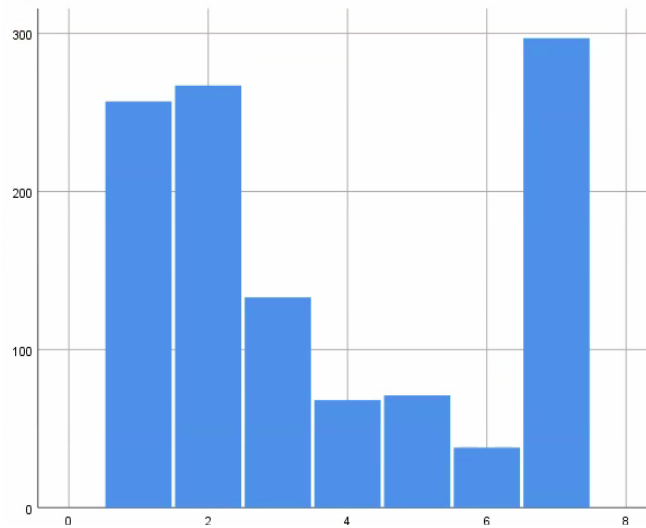
drie items antwoordoptie 8 “niet van toepassing (heb deze persoon niet)” gekozen, waardoor 7 als maximum wordt gezien. Zie figuren 1, 2 en 3 voor de verdelingen van de items.

Syntax:

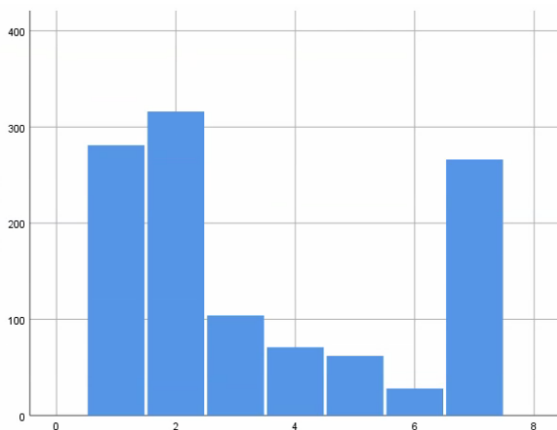
```
GRAPH
/HISTOGRAM=w1scb18b.
```

```
GRAPH
/HISTOGRAM=w1scb18c.
```

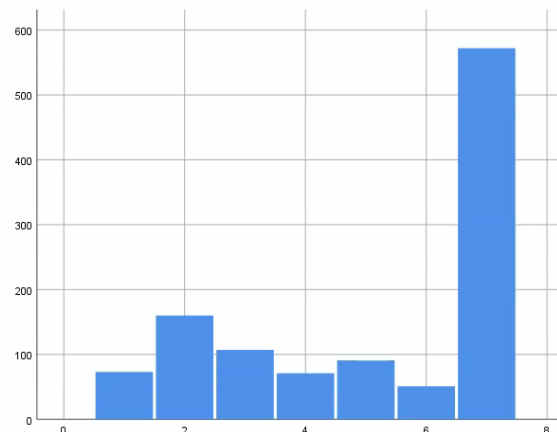
```
GRAPH
/HISTOGRAM=w1scb18d.
```



Figuur 1: Freq. verdeling contact met Marokkaanse buurtgenoten



Figuur 2: Freq. verdeling contact met Turkse buurtgenoten



Figuur 3: Freq. verdeling contact met Surinaamse/Antilliaanse buurtgenoten

Te zien is dat de groep respondenten die nooit contact heeft met buurtgenoten het grootst is wanneer deze buurtgenoten van Marokkaanse of Surinaamse/Antilliaanse komaf zijn. Daarnaast valt op dat de verdelingen van figuur 1 en 2 relatief veel op elkaar lijken, terwijl de verdeling van figuur 3 er niet op lijkt.

Om tot een bruikbare variabele te komen zijn de items allereerst gespiegeld, zodat een hogere waarde intensiever contact betekent. Vervolgens zijn er variabelen gemaakt die per etnische groep aangeven hoeveel contact zij hebben met co-etnische buurtgenoten, dus hoeveel contact een Marokkaanse respondent heeft met Marokkaanse buurtgenoten. Deze

variabelen zijn op hun beurt weer samengenomen tot één variabele die co-etnische sociale netwerken in de wijk meet. Dit gaf de volgende syntax en tabel 3:

```
RECODE w1scb18b w1scb18c w1scb18d (7=0) (6=1) (5=2) (4=3) (3=4) (2=5) (1=6) INTO
  SN_gespiegeld_turks Sn_gespiegeld_marok SN_gespiegeld_sur_antil.
VARIABLE LABELS SN_gespiegeld_turks 'Hogere waarde meer contact' /Sn_gespiegeld_marok
'Hogere '+
  'waarde meer contact' /SN_gespiegeld_sur_antil 'Hogere waarde is meer contact'.
EXECUTE.
```

```
IF(Ethnic_final = 0) SN_coeth_Marok = SN_gespiegeld_marok.
EXECUTE.
```

```
IF(Ethnic_final = 1) SN_coeth_Turks = SN_gespiegeld_turks.
EXECUTE.
```

```
IF(Ethnic_final = 2) SN_coeth_Sur_Antil = SN_gespiegeld_sur_antil.
EXECUTE.
```

Tabel 3: Freq. verdeling contact met co-etnische buurtgenoten

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 Vrijwel nooit	108	8,6	9,9	9,9
	1,00 1 keer per jaar	9	,7	,8	10,7
	2,00 Aantal keer per jaar	34	2,7	3,1	13,8
	3,00 1 keer per maand	55	4,4	5,0	18,9
	4,00 Aantal keer per maand	98	7,8	9,0	27,8
	5,00 1 of meerdere keren per week	361	28,8	33,1	60,9
	6,00 (bijna) Elke dag	427	34,1	39,1	100,0
	Total	1092	87,1	100,0	
Missing	System	162	12,9		
	Total	1254	100,0		

```
COMPUTE SN_co_ethn=0.
IF(Ethnic_final = 0) SN_co_ethn = SN_coeth_Marok.
IF(Ethnic_final = 1) SN_co_ethn = SN_coeth_Turks.
IF(Ethnic_final = 2) SN_co_ethn = SN_coeth_Sur_Antil.
EXECUTE.
```

```
FREQUENCIES VARIABLES=SN_co_ethn
/ORDER=ANALYSIS.
```

Omdat de middelste groepen klein zijn en er meerdere extreme waarden kwamen uit de logistische regressieanalyse behorend bij deze variabele, is ervoor gekozen categorieën samen te nemen. Dit resulteerde in de uiteindelijke variabele en diens verdeling:

Tabel 4: Freq. verdeling contact met co-etnische buurtgenoten

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 Vrijwel nooit	117	9,3	10,7	10,7
	1,00 Geregeld maar <1x per week	187	14,9	17,1	27,8
	2,00 Minstens 1x per week	788	62,8	72,2	100,0
	Total	1092	87,1	100,0	
Missing	System	162	12,9		
	Total	1254	100,0		

```
RECODE SN_co_ethn (0=0) (1=0) (2=1) (3=1) (4=1)
(5=2) (6=2) INTO New_SN_co_ethn.
VARIABLE LABELS New_SN_co_ethn 'Nieuwe
onafhankelijke variabele'.
EXECUTE.
```

```
FREQUENCIES VARIABLES=New_SN_co_ethn
/ORDER=ANALYSIS.
```

Opvallend is dat het merendeel van de respondenten minstens 1x per week contact heeft met co-etnische buurtgenoten en maar een relatief klein deel vrijwel nooit of geregeld.

Generatie migranten (moderator 1)

Deze variabele is ontstaan door de generaties van de drie etnische groepen samen te voegen en hier een dummy van te maken. De zelfgerapporteerde etniciteit van de (Turkse en Marokkaanse) respondenten en de twee eerder gemaakte dummy variabelen (generatie Surinamers en generatie Antillianen) zijn hiervoor gebruikt. Deze twee dummy variabelen hebben de volgende verdelingen:

FREQUENCIES VARIABLES=Antilliaans_1egen Antilliaans_2egen Surinaams_1egen Surinaams_2egen.
/ORDER=ANALYSIS.

Tabellen 5 en 6: Freq. verdeling Antilliaanse generaties Tabellen 7 en 8: Freq. verdeling Surinaamse generaties

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Niet van toepassing	1248	99,5	99,8	99,8
	Van toepassing	2	,2	,2	100,0
	Total	1250	99,7	100,0	
Missing	System	4	,3		
Total		1254	100,0		

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Niet van toepassing	1239	98,8	99,1	99,1
	Van toepassing	11	,9	,9	100,0
	Total	1250	99,7	100,0	
Missing	System	4	,3		
Total		1254	100,0		

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Niet van toepassing	1246	99,4	99,7	99,7
	Van toepassing	4	,3	,3	100,0
	Total	1250	99,7	100,0	
Missing	System	4	,3		
Total		1254	100,0		

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Niet van toepassing	1227	97,8	98,2	98,2
	Van toepassing	23	1,8	1,8	100,0
	Total	1250	99,7	100,0	
Missing	System	4	,3		
Total		1254	100,0		

Het samenvoegen van alle 1e generaties resulteerde in de uiteindelijke dummy variabele waarbij 1 betekent dat de respondent 1^e generatie migrant is en 0 2^e generatie migrant. Zie hiervoor onderstaande syntax en tabel:

Tabel 9: Freq. verdeling generatie

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00 2e generatie	217	17,3	17,4	17,4
	1,00 1e generatie	1033	82,4	82,6	100,0
	Total	1250	99,7	100,0	
Missing	System	4	,3		
Total		1254	100,0		

```
COMPUTE
Generatie=ANY(w1cethnic, 1, 3)
OR (Antilliaans_1egen = 1)
OR (Surinaams_1egen = 1).
EXECUTE.
```

```
FREQUENCIES VARIABLES=Generatie
/ORDER=ANALYSIS.
```

Te zien is dat de overgrote meerderheid tot de 1^e generatie behoort. Een zeer klein aandeel hiervan zijn Surinamers en Antillianen.

Etniciteit (moderator 2)

Deze moderator is gemeten aan de hand van dezelfde etniciteit variabele als die gebruikt is voor het creëren van de onafhankelijke variabele (co-etnische sociale netwerken in de wijk). Maar, van deze variabele zijn twee dummy's gemaakt, zodat zij opgenomen konden worden in de regressieanalyse. Dit leverde de volgende syntax en verdelingen op:

```
RECODE Ethnic_final (0=1) (1=0) (2=0) INTO d1_ethnic_Marokkaans.
VARIABLE LABELS d1_ethnic_Marokkaans 'Wel of niet Marokkaans'.
EXECUTE.
```

```
RECODE Ethnic_final (0=0) (1=1) (2=0) INTO d2_ethnic_Turks.
VARIABLE LABELS d2_ethnic_Turks 'Wel of niet Turks'.
EXECUTE.
```

Tabel 10: Frequentieverdeling dummy's etniciteit

d1_ethnic_Marokkaans Wel of niet Marokkaans

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 Niet Marokkaans	657	52,4	52,4	52,4
	1,00 Marokkaans	597	47,6	47,6	100,0
Total		1254	100,0	100,0	

d2_ethnic_Turks Wel of niet Turks

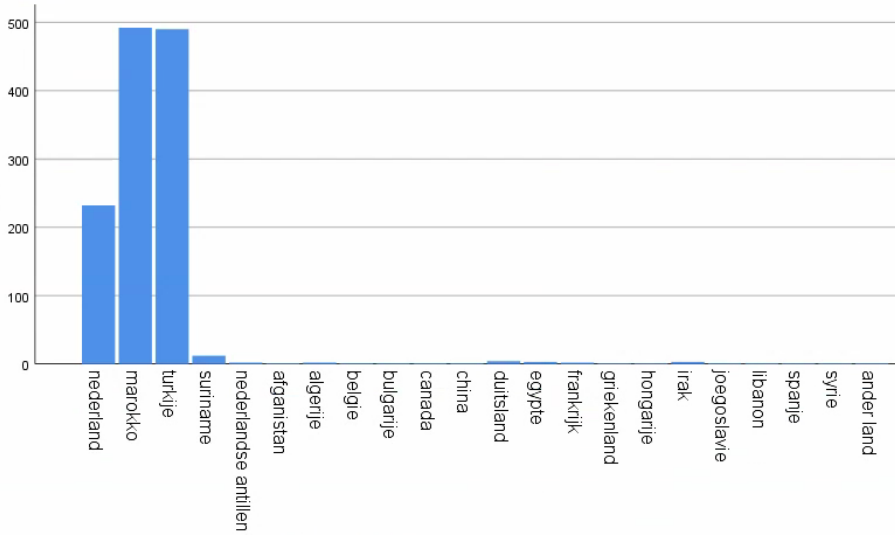
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 Niet Turks	624	49,8	49,8	49,8
	1,00 Turks	630	50,2	50,2	100,0
Total		1254	100,0	100,0	

```
FREQUENCIES VARIABLES=d1_ethnic_Marokkaans
d2_ethnic_Turks
/ORDER=ANALYSIS.
```

Te zien is dat de meerderheid Turks is en maar een klein percentage Surinaams/Antilliaans (2.2%).

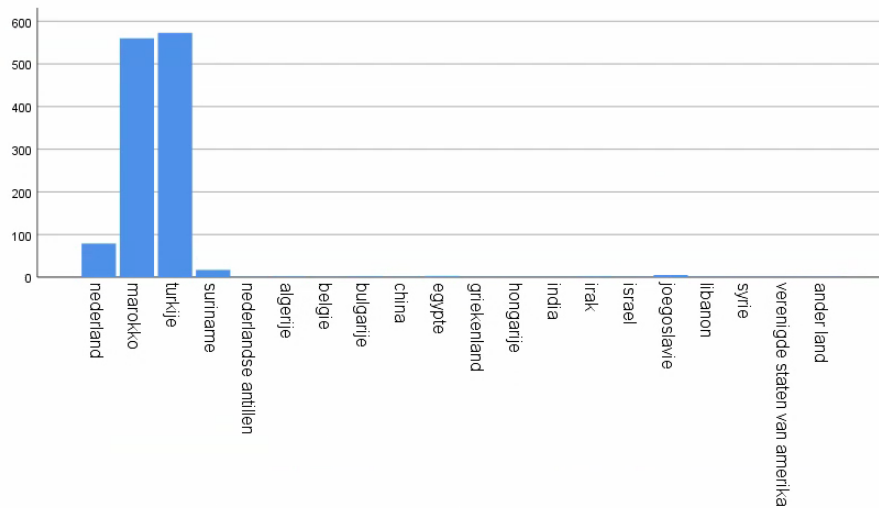
Keuze voor partner zelfde etniciteit (afhankelijke variabele)

Deze variabele is ontstaan door de etniciteit van de respondenten te vergelijken met de etniciteit van hun partner. Hiervoor is de eerder gebruikte etniciteit variabele weer gebruikt (Ethnic_final) en is er gekeken naar het geboorteland van de partner en diens ouders. Figuur 4, 5 en 6 laten de verdeling van de geboortelands zien van alle partijen.



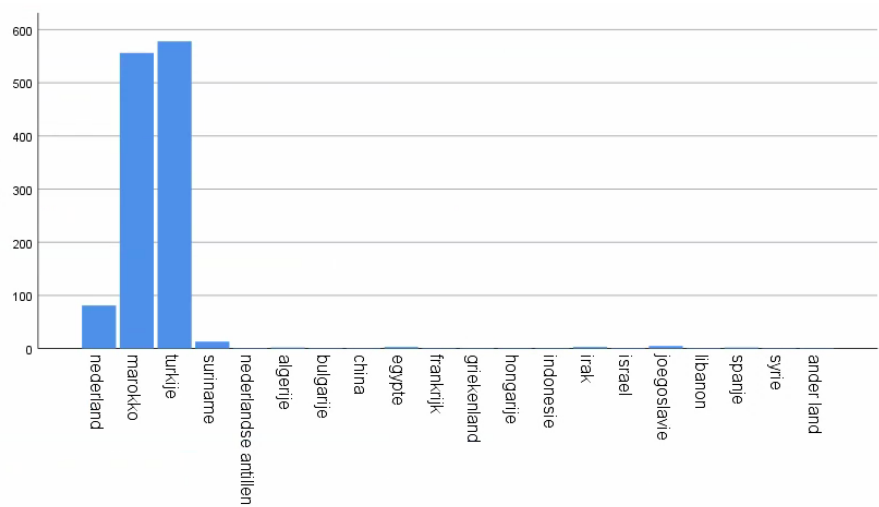
GRAPH
/BAR(SIMPLE)=COUNT BY w1fb17a.

Figuur 4: Frequentieverdeling geboorteland partner



GRAPH
/BAR(SIMPLE)=COUNT BY w1fb17b.

Figuur 5: Frequentieverdeling geboorteland vader van partner



GRAPH
/BAR(SIMPLE)=COUNT BY w1fb17c.

Figuur 6: Frequentieverdeling geboorteland moeder van partner

De verdelingen zien er ongeveer gelijk uit, met als grootste vier geboortelanden Nederland, Marokko, Turkije en Suriname.

Om de etniciteit van de partner te kunnen vergelijken met de etniciteit van de respondenten is het belangrijk de schalen gelijk te maken. Voor het vaststellen van de etniciteit van de partner is de eerdergenoemde definitie van etniciteit gebruikt; de definitie die door de NeLLS studie gehanteerd is. De volgende syntax en output hoort hierbij:

```
Compute Partner_ethnic = w1fb17a.
EXECUTE.
```

```
IF (w1fb17a ~= w1fb17c) Partner_ethnic = w1fb17c.
EXECUTE.
```

```
IF (w1fb17a=1 AND w1fb17c=1) Partner_ethnic = w1fb17b.
EXECUTE.
```

```
RECODE Partner_ethnic (16=2) (2=0) (3=1) (4=2) (5=2) (ELSE=Copy) INTO Partner_ethnic_new.
VARIABLE LABELS Partner_ethnic_new 'Nu overeenkomen met ethnic resp'.
EXECUTE.
```

```
COMPUTE Co_eth_huwelijk=Ethnic_final = Partner_ethnic_new.
EXECUTE.
```

Tabel 11: Frequentieverdeling wel/niet co-etnisch huwelijk

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 Niet van toepassing	71	5,7	5,7	5,7
	1.00 Wel van toepassing	1183	94,3	94,3	100,0
	Total	1254	100,0	100,0	

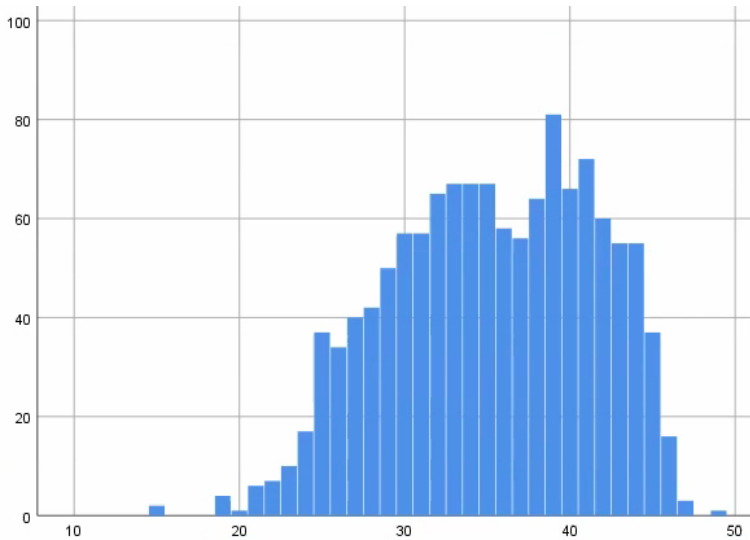
```
FREQUENCIES
VARIABLES=Co_eth_huwelijk
/ORDER=ANALYSIS.
```

Te zien is dat slechts 71 respondenten zich in een interetnisch huwelijk bevinden, wat neer komt op 5.7% van de totale groep.

Leeftijd (controlevariabele)

Deze variabele is niet bewerkt en is rechtstreeks opgenomen in de analyse. De onderstaande syntax, verdeling en descriptieve statistieken laten informatie zien over de respondenten hun leeftijd.

```
FREQUENCIES VARIABLES=w1cage
/FORMAT=NOTABLE
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
/HISTOGRAM
/ORDER=ANALYSIS.
```



Figuur 7: Freq. verdeling leeftijd ten tijde van interview

Te zien is dat mensen ouder dan 49 jaar geen onderdeel uitmaken van de steekproef. De verdeling is relatief normaal met één piek rond de 39.

Sekse (controlevariabele)

Van deze variabele is een dummy gemaakt, zodat de variabele opgenomen kon worden in de regressieanalyse. Dit leverde de volgende syntax en output op:

```
RECODE w1csex (1=0) (2=1) INTO w1csex_new.
VARIABLE LABELS w1csex_new 'Dummie sekse'.
EXECUTE.
```

Tabel 13: Frequentieverdeling sekse

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 Man	585	46,7	46,7	46,7
	1.00 Vrouw	669	53,3	53,3	100,0
	Total	1254	100,0	100,0	

Tabel 12: Descriptieve statistieken leeftijd

N	Valid	1254
	Missing	0
Mean		35,24
Std. Deviation		6,159
Minimum		15
Maximum		49

```
FREQUENCIES
VARIABLES=w1csex_new
/ORDER=ANALYSIS.
```

Te zien is dat de vrouwen iets in de meerderheid zijn, maar het verschil is klein.

Opleidingsniveau (controlevariabele)

Deze variabele is ontstaan door het samenvoegen van meerdere items. 14 Items bestaande uit de antwoordmogelijkheden "ja" en "nee" zijn samengevoegd om te kunnen meten wat iemands (hoogste behaalde) opleidingsniveau is. Voor elk item werd de vraag gesteld: "Hebt u deze opleiding met een diploma afgerond?" en werd per item een verschillend niveau

bevraagd, onder andere de lagere school, havo, mbo, hbo en de universiteit. De syntax en verdelingen van de niveaus zijn hieronder weergegeven.

Tabellen 14-27: frequentieverdelingen wel/niet diploma

**w1fa23b02 heeft u deze opleiding met een diploma afgerond?
lagere school**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	834	66,5	89,2	89,2
	2 nee	101	8,1	10,8	100,0
	Total	935	74,6	100,0	
Missing	System	319	25,4		
Total		1254	100,0		

**w1fa23b04 heeft u deze opleiding met een diploma afgerond?
mavo, vmbo-tl**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	244	19,5	80,8	80,8
	2 nee	58	4,6	19,2	100,0
	Total	302	24,1	100,0	
Missing	System	952	75,9		
Total		1254	100,0		

**w1fa23b03 heeft u deze opleiding met een diploma afgerond?
lbo, vmbo-kb/bbl**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	200	15,9	73,5	73,5
	2 nee	72	5,7	26,5	100,0
	Total	272	21,7	100,0	
Missing	System	982	78,3		
Total		1254	100,0		

**w1fa23b05 heeft u deze opleiding met een diploma afgerond?
havo**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	128	10,2	79,0	79,0
	2 nee	34	2,7	21,0	100,0
	Total	162	12,9	100,0	
Missing	System	1092	87,1		
Total		1254	100,0		

**w1fa23b06 heeft u deze opleiding met een diploma afgerond?
vwo/gymnasium**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	84	6,7	80,8	80,8
	2 nee	20	1,6	19,2	100,0
	Total	104	8,3	100,0	
Missing	System	1150	91,7		
Total		1254	100,0		

**w1fa23b08 heeft u deze opleiding met een diploma afgerond?
mbo-tussen/lang (mbo), secundai**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	253	20,2	85,8	85,8
	2 nee	42	3,3	14,2	100,0
	Total	295	23,5	100,0	
Missing	System	959	76,5		
Total		1254	100,0		

**w1fa23b07 heeft u deze opleiding met een diploma afgerond?
mbo-kort (kmbo), primair leerli**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	121	9,6	72,0	72,0
	2 nee	47	3,7	28,0	100,0
	Total	168	13,4	100,0	
Missing	System	1086	86,6		
Total		1254	100,0		

**w1fa23b09 heeft u deze opleiding met een diploma afgerond?
hbo**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	135	10,8	78,0	78,0
	2 nee	38	3,0	22,0	100,0
	Total	173	13,8	100,0	
Missing	System	1081	86,2		
Total		1254	100,0		

**w1fa23b10 heeft u deze opleiding met een diploma afgerond?
universiteit (bachelor)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	40	3,2	61,5	61,5
	2 nee	25	2,0	38,5	100,0
	Total	65	5,2	100,0	
Missing	System	1189	94,8		
Total		1254	100,0		

**w1fa23b12 heeft u deze opleiding met een diploma afgerond?
promotietraject**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	4	,3	80,0	80,0
	2 nee	1	,1	20,0	100,0
	Total	5	,4	100,0	
Missing	System	1249	99,6		
Total		1254	100,0		

**w1fa23b11 heeft u deze opleiding met een diploma afgerond?
universiteit (master, doctoraal)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	35	2,8	87,5	87,5
	2 nee	5	,4	12,5	100,0
	Total	40	3,2	100,0	
Missing	System	1214	96,8		
Total		1254	100,0		

**w1fa23b13 heeft u deze opleiding met een diploma afgerond?
buitenlandse opleiding, niet go**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	25	2,0	71,4	71,4
	2 nee	10	,8	28,6	100,0
	Total	35	2,8	100,0	
Missing	System	1219	97,2		
Total		1254	100,0		

**w1fa23b14 heeft u deze opleiding met een diploma afgerond?
buitenlandse opleiding, niet go**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	38	3,0	70,4	70,4
	2 nee	16	1,3	29,6	100,0
	Total	54	4,3	100,0	
Missing	System	1200	95,7		
Total		1254	100,0		

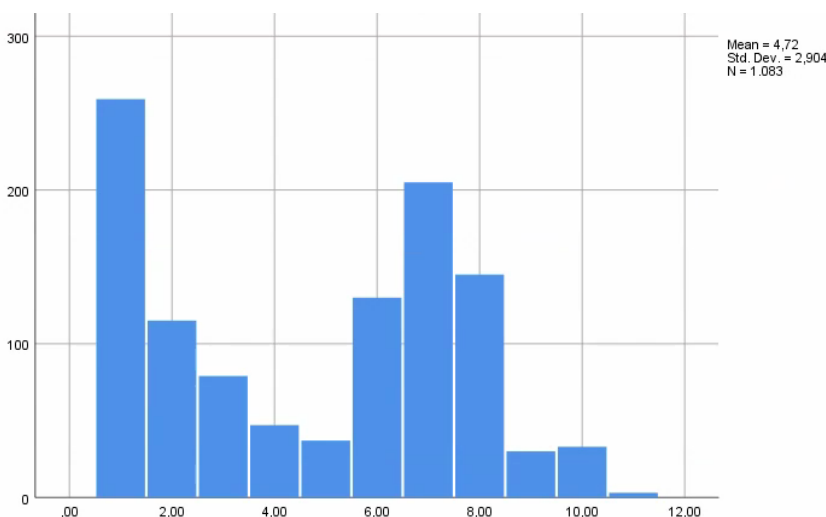
**w1fa23b15 heeft u deze opleiding met een diploma afgerond?
buitenlandse opleiding, niet go**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ja	25	2,0	80,6	80,6
	2 nee	6	,5	19,4	100,0
	Total	31	2,5	100,0	
Missing	System	1223	97,5		
Total		1254	100,0		

```
FREQUENCIES VARIABLES=w1fa23b02 w1fa23b03
w1fa23b04 w1fa23b05 w1fa23b06 w1fa23b07
w1fa23b08 w1fa23b09 w1fa23b10 w1fa23b11
w1fa23b12 w1fa23b13 w1fa23b14 w1fa23b15
/ORDER=ANALYSIS.
```

Te zien is dat er veel missende waarden zijn; naarmate het niveau omhooggaat lijken er ook meer missende waarden te zijn. Het kan zijn dat respondenten de vraag niet snapt, niet ingevuld hebben omdat zij de betreffende opleiding niet gevolgd hebben of niet wilden beantwoorden. Net als inkomen kan opleidingsniveau wellicht ook een gevoelig onderwerp zijn waar niet iedereen vrijuit over wil spreken. Bang voor stigmatisatie zou één van de redenen kunnen zijn.

De onderstaande bewerkingen hebben ertoe geleid dat van de 14 items één continue variabele gemaakt kon worden. Drie items zijn ondergebracht onder dezelfde waarden als drie andere items, namelijk "buitenlandse opleiding, niet goed in te delen, lager onderwijs", "buitenlandse opleiding, niet goed in te delen, middelbaar onderwijs" en "buitenlandse opleiding, niet goed in te delen, hoger onderwijs." Deze zijn ondergebracht onder dezelfde nieuwe waarden van respectievelijk "lagere school", "mbo-kort..." en "hbo." Zie figuur 8 voor de verdeling van deze nieuwe variabele.



Figuur 8: Frequentieverdeling opleidingsniveau

```
IF(w1fa23b02=1) Opleidingsniveau=1.
IF(w1fa23b03=1) Opleidingsniveau=2.
IF(w1fa23b04=1) Opleidingsniveau=3.
IF(w1fa23b05=1) Opleidingsniveau=4.
IF(w1fa23b06=1) Opleidingsniveau=5.
IF(w1fa23b07=1) Opleidingsniveau=6.
IF(w1fa23b08=1) Opleidingsniveau=7.
IF(w1fa23b09=1) Opleidingsniveau=8.
IF(w1fa23b10=1) Opleidingsniveau=9.
IF(w1fa23b11=1) Opleidingsniveau=10.
IF(w1fa23b12=1) Opleidingsniveau=11.
IF(w1fa23b13=1) Opleidingsniveau=1.
IF(w1fa23b14=1) Opleidingsniveau=6.
IF(w1fa23b15=1) Opleidingsniveau=8.
EXECUTE.
```

```
FREQUENCIES VARIABLES=Opleidingsniveau
/FORMAT=NOTABLE
/HISTOGRAM
/ORDER=ANALYSIS.
```

Te zien is dat een grote groep respondenten alleen de lagere school heeft afgerond met een diploma. Daarnaast is er ook een relatief grote groep die een mbo-diploma heeft. Aanzienlijk minder respondenten hebben een hbo of universitair diploma.

Bijlage 2: Bespreking statistieken en analyses

In deze bijlage zal de logistische regressieanalyse die uitgevoerd is toegelicht worden. De syntax en SPSS-output behorend bij de univariate en bivariate descriptieve statistieken zullen eerst besproken worden.

Univariate descriptieve statistieken

De univariate statistieken bestaan uit gemiddelden, standaarddeviaties, minimum- en maximumwaarden en frequenties uitgedrukt in percentages en absolute aantallen. Dit hangt af van het soort variabele. Zo is zijn voor de continue variabele *leeftijd* en *opleidingsniveau* de volgende output verkregen:

Tabel 1: univariate statistieken leeftijd en opleidingsniveau

	N	Minimum	Maximum	Mean	Std. Deviation
w1 cage age at time of interview	1254	15	49	35,24	6,159
Opleidingsniveau	1083	1,00	11,00	4,7156	2,90363
Valid N (listwise)	1083				

```
DESCRIPTIVES VARIABLES=w1cage  
Opleidingsniveau  
/STATISTICS=MEAN STDDEV MIN MAX.
```

Zie bijlage 1 figuur 7 en 8 voor de verdeling van respectievelijk *leeftijd* en *opleidingsniveau*. Te zien is dat de gemiddelde leeftijd van 35 jaar relatief laag is, wat verklaard kan worden door de maximumleeftijd van slechts 49 jaar. Vanwege de vele categorieën is opleidingsniveau meegenomen als een continue variabele. Aan de relatief hoge standaarddeviatie van 2.9 is te zien dat er veel spreiding is in de respondenten hun opleidingsniveau.

De frequenties behorend bij de categorische variabelen zijn uitgedrukt in absolute aantallen en percentages. De valide percentages ("Valid Percent") zijn verwerkt in de resultatensectie van dit onderzoek. Middels de volgende syntax is de onderstaande output verkregen:

```
FREQUENCIES VARIABLES=New_SN_co_ethn Co_eth_huwelijk Ethnic_final Generatie w1csex_new  
/ORDER=ANALYSIS.
```

Tabellen 2-6: Frequentieverdelingen categorische variabelen

New_SN_co_ethn Nieuwe onafhankelijke variabele

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 Vrijwel nooit	117	9,3	10,7	10,7
	1,00 Geregeld maar <1x per week	187	14,9	17,1	27,8
	2,00 Minstens 1x per week	788	62,8	72,2	100,0
	Total	1092	87,1	100,0	
Missing	System	162	12,9		
Total		1254	100,0		

Co_eth_huwelijk Afhankelijke variabele

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 Niet van toepassing	71	5,7	5,7	5,7
	1,00 Wel van toepassing	1183	94,3	94,3	100,0
	Total	1254	100,0	100,0	

w1csex_new Dummie sekse

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 Man	585	46,7	46,7	46,7
	1,00 Vrouw	669	53,3	53,3	100,0
	Total	1254	100,0	100,0	

Ethnic_final Drie groepen; Marokkaans, Turks, Antilliaans+Surinaams

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 Marokkaans	597	47,6	47,6	47,6
	1,00 Turks	630	50,2	50,2	97,8
	2,00 Antilliaans/Surinaams	27	2,2	2,2	100,0
	Total	1254	100,0	100,0	

Generatie Generatie migranten

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 2e generatie	217	17,3	17,4	17,4
	1,00 1e generatie	1033	82,4	82,6	100,0
	Total	1250	99,7	100,0	
Missing	System	4	,3		
Total		1254	100,0		

Te zien is dat de verdelingen, op sekse na, scheef zijn. Een kleine groep respondenten (117 resp., 10.7%) heeft vrijwel nooit contact met co-etnische buurtgenoten en een kleine groep (71 resp., 5.7%) is getrouwd met iemand die een andere etniciteit heeft dan henzelf.

Bivariate descriptieve statistieken

De bivariate statistieken bestaan uit correlaties en gemiddelden tussen groepen. Correlaties tussen continue variabelen (leeftijd en opleidingsniveau) en dichotome variabelen (sekse, generatie, zelfde etniciteit resp. en partner) zijn berekend aan de hand van een t-toets die gemiddelde verschillen tussen de groepen toetst. Correlaties tussen twee dichotome variabelen zijn berekend aan de hand van Pearson's correlatie als indicatie voor de phi-coëfficiënt (ϕ). Correlaties tussen continue en categorische variabelen zijn berekend op basis van de ANOVA en de F-toets via de formule $R = \sqrt{1 - \frac{n-k}{n-k+(k-1)F}}$. Tot slot zijn correlaties tussen twee categorische variabelen berekend op basis van Cramer's V.

Voor de samenhang tussen de continue en dichotome variabelen zijn t-toetsen berekend die hieronder, met bijbehorende syntax, weergegeven zijn.

Tabel 7: T-toets samenhang leeftijd en sekse

Group Statistics					
	w1csex_new Dummie sekse	N	Mean	Std. Deviation	Std. Error Mean
w1 cage age at time of interview	,00 Man	585	36,54	5,705	,236
	1,00 Vrouw	669	34,10	6,319	,244

T-TEST GROUPS=w1csex_new(0 1)
/MISSING=ANALYSIS
/VARIABLES=w1cage
/CRITERIA=CI(.95).

Independent Samples Test										
		Levene's Test for Equality of Variances				t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
w1 cage age at time of interview	Equal variances assumed	12,697	,000	7,127	1252	,000	2,437	,342	1,766	3,108
	Equal variances not assumed			7,175	1250,722	,000	2,437	,340	1,771	3,103

Tabel 8: T-toets samenhang leeftijd en generatie

Group Statistics					
	Generatie migranten	N	Mean	Std. Deviation	Std. Error Mean
w1 cage age at time of interview	,00 2e generatie	217	30,69	5,287	,359
	1,00 1e generatie	1033	36,21	5,895	,183

T-TEST GROUPS=Generatie(0 1)
/MISSING=ANALYSIS
/VARIABLES=w1cage
/CRITERIA=CI(.95).

Independent Samples Test										
		Levene's Test for Equality of Variances				t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
w1 cage age at time of interview	Equal variances assumed	8,929	,003	-12,772	1248	,000	-5,526	,433	-6,375	-4,677
	Equal variances not assumed			-13,710	338,683	,000	-5,526	,403	-6,319	-4,733

Tabel 9: T-toets samenhang leeftijd en wel/niet co-etnisch huwelijk

Group Statistics					
	Co_eth_huwelijk Afhankelijke variabele	N	Mean	Std. Deviation	Std. Error Mean
w1 cage age at time of interview	,00 Niet van toepassing	71	35,89	5,437	,645
	1,00 Wel van toepassing	1183	35,20	6,200	,180

T-TEST GROUPS=Co_eth_huwelijk(0 1)
/MISSING=ANALYSIS
/VARIABLES=w1cage
/CRITERIA=CI(.95).

Independent Samples Test										
		Levene's Test for Equality of Variances				t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
w1 cage age at time of interview	Equal variances assumed	3,024	,082	,914	1252	,361	,688	,753	-,789	2,164
	Equal variances not assumed			1,027	81,324	,308	,688	,670	-,645	2,021

Tabel 10: T-toets samenhang opleidingsniveau en sekse

Group Statistics					
	w1csex_new Dummie sekse	N	Mean	Std. Deviation	Std. Error Mean
Opleidingsniveau	,00 Man	522	4,7261	2,93951	,12866
	1,00 Vrouw	561	4,7059	2,87243	,12127

T-TEST GROUPS=w1csex_new(0 1)
/MISSING=ANALYSIS
/VARIABLES=Opleidingsniveau
/CRITERIA=CI(.95).

Independent Samples Test										
		Levene's Test for Equality of Variances				t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Opleidingsniveau	Equal variances assumed	,656	,418	,114	1081	,909	,02017	,17666	-,32646	,36681
	Equal variances not assumed			,114	1071,294	,909	,02017	,17681	-,32675	,36710

Tabel 11: T-toets samenhang opleidingsniveau en generatie

Group Statistics					
	Generatie migranten	N	Mean	Std. Deviation	Std. Error Mean
Opleidingsniveau	,00 2e generatie	208	5,7115	2,56386	,17777
	1,00 1e generatie	871	4,4696	2,93041	,09929

T-TEST GROUPS=Generatie(0 1)
/MISSING=ANALYSIS
/VARIABLES=Opleidingsniveau
/CRITERIA=CI(.95).

Independent Samples Test										
		Levene's Test for Equality of Variances				t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Opleidingsniveau	Equal variances assumed	33,688	,000	5,620	1077	,000	1,24196	,22100	,80833	1,67559
	Equal variances not assumed			6,099	348,238	,000	1,24196	,20362	,84148	1,64245

Tabel 12: T-toets samenhang opleidingsniveau en wel/niet co-etnisch huwelijk

Group Statistics					
	Co_eth_huwelijk Afhankelijke variabele	N	Mean	Std. Deviation	Std. Error Mean
Opleidingsniveau	,00 Niet van toepassing	65	5,9846	2,89662	,35928
	1,00 Wel van toepassing	1018	4,6346	2,88659	,09047

T-TEST GROUPS=Co_eth_huwelijk(0 1)
/MISSING=ANALYSIS
/VARIABLES=Opleidingsniveau
/CRITERIA=CI(.95).

Independent Samples Test										
		Levene's Test for Equality of Variances				t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Opleidingsniveau	Equal variances assumed	1,438	,231	3,655	1081	,000	1,35004	,36937	,62528	2,07480
	Equal variances not assumed			3,644	72,355	,001	1,35004	,37050	,61153	2,08855

Te zien is dat er significante verschillen zijn tussen het opleidingsniveau van de 1^e generatie en de 2^e generatie; gemiddeld genomen is de 2^e generatie hoger opgeleid dan de 1^e generatie. Daarnaast zijn er significante verschillen tussen respondenten die zich in een co-etnisch huwelijk bevinden en zij die dat niet zijn, wat betreft opleidingsniveau; gemiddeld

genomen zijn degenen die zich in een interetnisch huwelijk bevinden iets hoger opgeleid dan degenen die zich in een co-etnisch huwelijk bevinden.

Voor de samenhang tussen de twee continue variabelen en tussen de dichotome variabelen onderling zijn Pearson's correlatie (als indicatie voor de phi-coëfficiënt ϕ) berekend en hieronder weergegeven.

Tabel 13: Correlatie tussen opleidingsniveau en leeftijd

		w1cage age at time of interview	Opleidingsniv eau
w1cage age at time of interview	Pearson Correlation	1	-,131**
	Sig. (2-tailed)		,000
	N	1254	1083
Opleidingsniveau	Pearson Correlation	-,131**	1
	Sig. (2-tailed)	,000	
	N	1083	1083

** . Correlation is significant at the 0.01 level (2-tailed).

```
CORRELATIONS
/VARIABLES=w1cage Opleidingsniveau
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

Tabel 14: Correlatie tussen sekse, generatie en wel/niet co-etnisch huwelijk

		w1csex_new Dummie sekse	Generatie Generatie migranten	Co_eth_huwe lijk Afhankelijke variabele
w1csex_new Dummie sekse	Pearson Correlation	1	-,069*	-,001
	Sig. (2-tailed)		,015	,976
	N	1254	1250	1254
Generatie Generatie migranten	Pearson Correlation	-,069*	1	,118**
	Sig. (2-tailed)	,015		,000
	N	1250	1250	1250
Co_eth_huwelijk Afhankelijke variabele	Pearson Correlation	-,001	,118**	1
	Sig. (2-tailed)	,976	,000	
	N	1254	1250	1254

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

```
CORRELATIONS
/VARIABLES=w1csex_new Generatie
Co_eth_huwelijk
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

Te zien is dat er alleen geen significant verband lijkt te zijn tussen iemands sekse en of diegene zich al dan niet in een co-etnisch huwelijk bevindt. Het positieve verband van generatie en co-etnisch huwelijk laat daarnaast zien dat de 1^e generatie meer kans heeft om zich in een co-etnisch huwelijk te bevinden van de 2^e generatie.

Voor de samenhang tussen de continue en categorische variabelen is er gekeken naar de

ANOVA en de F-toets. Correlaties zijn berekend met de formule $R = \sqrt{1 - \frac{n-k}{n-k+(k-1)F}}$. Op

basis van Fisher's Z-transformatie kan gecontroleerd worden of de correlaties significant zijn

of niet. Wanneer de correlaties groter zijn, in absolute waarde, dan $\frac{a-1}{a+1}$ met $a = e^{\left(\frac{2 \times 1.96}{\sqrt{n-3}}\right)}$, dan zijn de correlaties significant. De onderstaande output en berekeningen laten zien hoe de resultaten tot stand zijn gekomen.

Tabel 15: ANOVA behorend bij leeftijd en etniciteit

ANOVA

w1cage age at time of interview

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	267,766	2	133,883	3,543	,029
Within Groups	47269,941	1251	37,786		
Total	47537,707	1253			

ONEWAY w1cage BY Ethnic_final /MISSING ANALYSIS.

De berekening $\sqrt{1 - \frac{1254-3}{1254-3+(3-1)3.543}}$ geeft een R waarde van 0.08

$e^{\left(\frac{2 \times 1.96}{\sqrt{1254-3}}\right)} = a = 1.117$. $\frac{0.117}{2.117}$ geeft een drempelwaarde van 0.055, wat betekent dat de correlatie significant is.

Tabel 16: ANOVA behorend bij leeftijd en co-etnisch SN

w1cage age at time of interview

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	30,738	2	15,369	,410	,664
Within Groups	40820,723	1089	37,485		
Total	40851,461	1091			

ONEWAY w1cage BY New_SN_co_ethn /MISSING ANALYSIS.

De berekening $\sqrt{1 - \frac{1092-3}{1092-3+(3-1)0.410}}$ geeft een R waarde van 0.03.

De drempelwaarde van 0.059 laat zien dat de correlatie niet significant is.

Tabel 17: ANOVA behorend bij opleidingsniveau en etniciteit

ANOVA

Opleidingsniveau

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	52,880	2	26,440	3,148	,043
Within Groups	9069,526	1080	8,398		
Total	9122,406	1082			

ONEWAY Opleidingsniveau BY Ethnic_final /MISSING ANALYSIS.

De berekening $\sqrt{1 - \frac{1083-3}{1083-3+(3-1)3.148}}$ geeft een R waarde van 0.08.

De drempelwaarde van 0.06 laat zien dat de correlatie significant is.

Tabel 18: ANOVA behorend bij opleidingsniveau en co-etnisch SN

ANOVA					
Opleidingsniveau	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	195,786	2	97,893	11,998	,000
Within Groups	7784,038	954	8,159		
Total	7979,824	956			

ONEWAY Opleidingsniveau BY New_SN_co_ethn
/MISSING ANALYSIS.

De berekening $\sqrt{1 - \frac{957-3}{957-3+(3-1)11.998}}$ geeft een R waarde van 0.16.

De drempelwaarde van 0.063 laat zien dat de correlatie significant is.

Te zien is dat de F-toetsen significant zijn, behalve bij leeftijd en co-etnisch contact in de wijk. Dit betekent dat er significante verschillen zijn tussen de etnische groepen wat betreft opleidingsniveau en leeftijd. Daarnaast zijn er verschillen tussen respondenten die nooit, geregeld of wekelijks contact hebben met co-etnische buurtgenoten wat betreft hun opleidingsniveau.

Tot slot is de samenhang tussen twee categorische variabelen berekend op basis van Cramer's V. Dit leverde de volgende syntax en output op:

Tabel 19: Cramer's V etniciteit en co-etnisch SN

Symmetric Measures			
		Value	Approximate Significance
Nominal by Nominal	Phi	,156	,000
	Cramer's V	,110	,000
N of Valid Cases		1092	

CROSSTABS
/TABLES=Ethnic_final BY New_SN_co_ethn
/FORMAT=AVALUE TABLES
/STATISTICS=PHI
/CELLS=COUNT
/COUNT ROUND CELL.

Tabel 20: Cramer's V sekse en co-etnisch SN

Symmetric Measures			
		Value	Approximate Significance
Nominal by Nominal	Phi	,095	,007
	Cramer's V	,095	,007
N of Valid Cases		1092	

CROSSTABS
/TABLES=w1csex_new BY New_SN_co_ethn
/FORMAT=AVALUE TABLES
/STATISTICS=PHI
/CELLS=COUNT
/COUNT ROUND CELL.

Tabel 21: Cramer's V sekse en etniciteit

Symmetric Measures			
		Value	Approximate Significance
Nominal by Nominal	Phi	,043	,321
	Cramer's V	,043	,321
N of Valid Cases		1254	

```
CROSSTABS
/TABLES=w1csex BY Ethnic_final
/FORMAT=AVALUE TABLES
/STATISTICS=PHI
/CELLS=COUNT
/COUNT ROUND CELL.
```

Tabel 22: Cramer's V etniciteit en generatie

Symmetric Measures			
		Value	Approximate Significance
Nominal by Nominal	Phi	,096	,003
	Cramer's V	,096	,003
N of Valid Cases		1250	

```
CROSSTABS
/TABLES=Generatie BY Ethnic_final
/FORMAT=AVALUE TABLES
/STATISTICS=PHI
/CELLS=COUNT
/COUNT ROUND CELL.
```

Tabel 23: Cramer's V generatie en co-etnisch SN

Symmetric Measures			
		Value	Approximate Significance
Nominal by Nominal	Phi	,107	,002
	Cramer's V	,107	,002
N of Valid Cases		1088	

```
CROSSTABS
/TABLES=Generatie BY New_SN_co_ethn
/FORMAT=AVALUE TABLES
/STATISTICS=PHI
/CELLS=COUNT
/COUNT ROUND CELL.
```

Tabel 24: Cramer's V co-etnisch SN en wel/niet co-etnisch huwelijk

Symmetric Measures			
		Value	Approximate Significance
Nominal by Nominal	Phi	,228	,000
	Cramer's V	,228	,000
N of Valid Cases		1092	

```
CROSSTABS
/TABLES=Co_eth_huwelijk BY New_SN_co_ethn
/FORMAT=AVALUE TABLES
/STATISTICS=PHI
/CELLS=COUNT
/COUNT ROUND CELL.
```

Tabel 25: Cramer's V etniciteit en wel/niet co-etnisch huwelijk

Symmetric Measures			
		Value	Approximate Significance
Nominal by Nominal	Phi	,341	,000
	Cramer's V	,341	,000
N of Valid Cases		1254	

```
CROSSTABS
/TABLES=Co_eth_huwelijk BY Ethnic_final
/FORMAT=AVALUE TABLES
/STATISTICS=PHI
/CELLS=COUNT
/COUNT ROUND CELL.
```

Te zien is dat, op etniciteit en sekse na, alle waarden significant zijn. Het lijkt erop dat de variabelen met elkaar samenhangen, met name etniciteit en het wel of niet co-etnisch getrouwd zijn lijkt het sterkst samen te hangen vergeleken met de anderen.

Logistische regressieanalyse

De logistische regressieanalyse die gebruikt is om de hypothesen te toetsen is hieronder weergegeven, met bijbehorende syntax. Op een hiërarchische wijze zijn variabelen aan de modellen toegevoegd; model 0 bevat enkel de afhankelijke variabele, 1 bevat daarnaast ook de controlevariabelen, aan model 2 is de onafhankelijke variabele toegevoegd, model 3 de moderatoren, model 4 de eerste interactieterm en model 5 bevat ook nog de tweede interactieterm. Omdat de missende waarden niet meegenomen zijn in de analyse is de steekproefgrootte gelijk voor elk model ($n=953$).

```
LOGISTIC REGRESSION VARIABLES Co_eth_huwelijk
/METHOD=ENTER w1cage w1csex_new Opleidingsniveau
/METHOD=ENTER w1cage w1csex_new Opleidingsniveau New_SN_co_ethn
/METHOD=ENTER w1cage w1csex_new Opleidingsniveau New_SN_co_ethn Generatie
d1_ethnic_Marokkaans
  d2_ethnic_Turks
/METHOD=ENTER w1cage w1csex_new Opleidingsniveau New_SN_co_ethn Generatie
d1_ethnic_Marokkaans
  d2_ethnic_Turks New_snXgeneratie
/METHOD=ENTER w1cage w1csex_new Opleidingsniveau New_SN_co_ethn Generatie
d1_ethnic_Marokkaans
  d2_ethnic_Turks New_snXgeneratie New_snXd1_Marok New_snXd2_Turks
/CONTRAST (w1csex_new)=Indicator(1)
/CONTRAST (New_SN_co_ethn)=Indicator(1)
/CONTRAST (Generatie)=Indicator(1)
/CONTRAST (d1_ethnic_Marokkaans)=Indicator(1)
/CONTRAST (d2_ethnic_Turks)=Indicator(1)
/CONTRAST (New_snXgeneratie)=Indicator(1)
/CONTRAST (New_snXd1_Marok)=Indicator(1)
/CONTRAST (New_snXd2_Turks)=Indicator(1)
/SAVE=COOK LEVER
/CLASSPLOT
/CASEWISE OUTLIER(2)
/PRINT=GOODFIT CI(95)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

Tabel 26: Overzicht geselecteerde en ongeselecteerde cases

Case Processing Summary			
Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	953	76,0
	Missing Cases	301	24,0
	Total	1254	100,0
Unselected Cases		0	,0
Total		1254	100,0

a. If weight is in effect, see classification table for the total number of cases.

Tabel 27: Model 0 logistische regressieanalyse

Variables in the Equation							
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	2,755	,137	406,723	1	,000	15,719

Variables not in the Equation							
		Score	df	Sig.			
Step 0	Variables	age at time of interview	,684	1	,408		
		Dummie sekse(1)	,080	1	,777		
		Opleidingsniveau	15,978	1	,000		
Overall Statistics		18,523	3	,000			

Tabel 28: Model 1 modelfit

Omnibus Tests of Model Coefficients				
		Chi-square	df	Sig.
Step 1	Step	18,958	3	,000
	Block	18,958	3	,000
	Model	18,958	3	,000

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	412,651 ^a	,020	,054

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	10,206	8	,251

Tabel 29: Model 1 logistische regressieanalyse

Variables in the Equation									
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	age at time of interview	-,036	,024	2,198	1	,138	,965	,921	1,012
	Dummie sekse(1)	-,206	,285	,523	1	,469	,814	,466	1,422
	Opleidingsniveau	-,213	,053	16,279	1	,000	,808	,729	,896
	Constant	5,279	1,005	27,591	1	,000	196,262		

a. Variable(s) entered on step 1: age at time of interview, Dummie sekse, Opleidingsniveau.

Tabel 30: Model 2 logistische regressieanalyse

Variables in the Equation									
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	age at time of interview	-,035	,025	2,004	1	,157	,965	,920	1,014
	Dummie sekse(1)	-,266	,297	,806	1	,369	,766	,428	1,370
	Opleidingsniveau	-,172	,054	10,113	1	,001	,842	,757	,936
	Nieuwe onafhankelijke variabele			35,482	2	,000			
	Nieuwe onafhankelijke variabele(1)	1,226	,386	10,108	1	,001	3,408	1,600	7,257
	Nieuwe onafhankelijke variabele(2)	1,941	,327	35,217	1	,000	6,963	3,668	13,217
	Constant	3,716	1,078	11,873	1	,001	41,091		

a. Variable(s) entered on step 1: age at time of interview, Dummie sekse, Opleidingsniveau, Nieuwe onafhankelijke variabele.

Tabel 31: Model 2 modelfit statistieken

Omnibus Tests of Model Coefficients				
		Chi-square	df	Sig.
Step 1	Step	32,807	2	,000
	Block	32,807	2	,000
	Model	51,765	5	,000

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	379,844 ^a	,053	,145

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	7,960	8	,437

Tabel 32: Model 3 modelfit statistieken

Omnibus Tests of Model Coefficients				
		Chi-square	df	Sig.
Step 1	Step	65,614	3	,000
	Block	65,614	3	,000
	Model	117,379	8	,000

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	314,230 ^a	,116	,318

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	3,443	8	,904

Tabel 33: Model 3 logistische regressieanalyse

Variables in the Equation									
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	age at time of interview	-,030	,029	1,075	1	,300	,970	,917	1,027
	Dummie sekse(1)	-,059	,329	,032	1	,858	,943	,495	1,796
	Opleidingsniveau	-,141	,060	5,543	1	,019	,869	,773	,977
	Nieuwe onafhankelijke variabele			20,608	2	,000			
	Nieuwe onafhankelijke variabele(1)	1,262	,437	8,334	1	,004	3,532	1,500	8,320
	Nieuwe onafhankelijke variabele(2)	1,614	,362	19,935	1	,000	5,024	2,474	10,205
	Generatie migranten(1)	,844	,364	5,380	1	,020	2,325	1,140	4,742
	Wel of niet Marokkaans (1)	2,435	,490	24,736	1	,000	11,412	4,372	29,790
	Wel of niet Turks(1)	4,084	,574	50,634	1	,000	59,381	19,280	182,884
Constant	,021	1,325	,000	1	,987	1,022			

a. Variable(s) entered on step 1: age at time of interview, Dummie sekse, Opleidingsniveau, Nieuwe onafhankelijke variabele, Generatie migranten, Wel of niet Marokkaans, Wel of niet Turks.

Tabel 34: Model 4 logistische regressieanalyse

		Variables in the Equation							
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	age at time of interview	-,027	,029	,849	1	,357	,973	,919	1,031
	Dummie sekse(1)	-,018	,334	,003	1	,958	,983	,510	1,892
	Opleidingsniveau	-,137	,060	5,197	1	,023	,872	,774	,981
	Nieuwe onafhankelijke variabele			9,513	2	,009			
	Nieuwe onafhankelijke variabele(1)	1,373	,724	3,599	1	,058	3,948	,956	16,314
	Nieuwe onafhankelijke variabele(2)	2,026	,671	9,130	1	,003	7,587	2,038	28,242
	Generatie migranten(1)	1,124	,594	3,582	1	,058	3,077	,961	9,854
	Wel of niet Marokkaans (1)	2,507	,503	24,852	1	,000	12,270	4,579	32,881
	Wel of niet Turks(1)	4,136	,583	50,404	1	,000	62,548	19,968	195,924
	New_snXgeneratie			,564	2	,754			
	New_snXgeneratie(1)	-,165	,912	,033	1	,857	,848	,142	5,067
	New_snXgeneratie(2)	-,593	,809	,536	1	,464	,553	,113	2,700
	Constant	-,366	1,432	,066	1	,798	,693		

a. Variable(s) entered on step 1: age at time of interview, Dummie sekse, Opleidingsniveau, Nieuwe onafhankelijke variabele, Generatie migranten, Wel of niet Marokkaans, Wel of niet Turks, New_snXgeneratie.

Tabel 35: Model 4 modelfit statistieken

Omnibus Tests of Model Coefficients				
		Chi-square	df	Sig.
Step 1	Step	,576	2	,750
	Block	,576	2	,750
	Model	117,955	10	,000

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	313,654 ^a	,116	,320

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	5,923	8	,656

Tabel 36: Model 5 modelfit statistieken

Omnibus Tests of Model Coefficients				
		Chi-square	df	Sig.
Step 1	Step	5,327	4	,255
	Block	5,327	4	,255
	Model	123,282	14	,000

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	308,327 ^a	,121	,333

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	2,090	8	,978

Tabel 37: Model 5 logistische regressieanalyse

		Variables in the Equation							95% C.I. for EXP(B)	
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper	
Step 1 ^a	age at time of interview	-,028	,029	,885	1	,347	,973	,918	1,030	
	Dummie sekse(1)	,010	,336	,001	1	,977	1,010	,523	1,950	
	Opleidingsniveau	-,145	,062	5,508	1	,019	,865	,767	,976	
	Nieuwe onafhankelijke variabele			5,655	2	,059				
	Nieuwe onafhankelijke variabele(1)	1,692	1,572	1,158	1	,282	5,429	,249	118,248	
	Nieuwe onafhankelijke variabele(2)	3,352	1,454	5,318	1	,021	28,560	1,654	493,190	
	Generatie migranten(1)	1,185	,604	3,847	1	,050	3,269	1,001	10,678	
	Wel of niet Marokkaans (1)	3,340	1,182	7,983	1	,005	28,215	2,782	286,199	
	Wel of niet Turks(1)	4,642	1,288	12,996	1	,000	103,711	8,315	1293,556	
	New_snXgeneratie			,639	2	,727				
	New_snXgeneratie(1)	-,043	,974	,002	1	,965	,958	,142	6,460	
	New_snXgeneratie(2)	-,589	,805	,536	1	,464	,555	,115	2,687	
	New_snXd1_Marok			1,180	2	,554				
	New_snXd1_Marok(1)	-,699	1,448	,233	1	,629	,497	,029	8,495	
	New_snXd1_Marok(2)	-1,488	1,423	1,094	1	,296	,226	,014	3,672	
	New_snXd2_Turks			,803	2	,669				
	New_snXd2_Turks(1)	16,980	4173,878	,000	1	,997	23686413,77	,000	.	
	New_snXd2_Turks(2)	-1,392	1,554	,803	1	,370	,248	,012	5,226	
	Constant	-1,048	1,770	,350	1	,554	,351			

a. Variable(s) entered on step 1: age at time of interview, Dummie sekse, Opleidingsniveau, Nieuwe onafhankelijke variabele, Generatie migranten, Wel of niet Marokkaans, Wel of niet Turks, New_snXgeneratie, New_snXd1_Marok, New_snXd2_Turks.

Omdat de interactieterm *New_snXd2_Turks(1)* een extreem grote standaarddeviatie en odds ratio heeft in model 5, is ervoor gekozen de referentiegroep van etniciteit (Surinaams/Antilliaans) te veranderen. Het aandeel Surinaamse/Antilliaanse respondenten in namelijk erg klein en de verwachting bestond dat wanneer de referentiegroep Marokkaans zou zijn, de extreme waarden zouden veranderen. Echter, zoals te zien is in de onderstaande tabel, is dit niet veranderd. Daarom is ervoor gekozen de Surinaamse/Antilliaanse groep als referentiegroep te behouden.

```
RECODE Ethnic_final (0=0) (1=0) (2=1) INTO d3_ethnic_suri_antil.
VARIABLE LABELS d3_ethnic_suri_antil 'Wel of niet Surinaams/Antilliaans'.
EXECUTE.
```

```
COMPUTE New_snXd3_suri_antil=New_SN_co_ethn * d3_ethnic_suri_antil.
EXECUTE.
```

```
LOGISTIC REGRESSION VARIABLES Co_eth_huwelijk
/METHOD=ENTER w1cage w1csex_new Opleidingsniveau New_SN_co_ethn Generatie
d3_ethnic_suri_antil d2_ethnic_Turks New_snXgeneratie New_snXd3_suri_antil New_snXd2_Turks
```

```

/CONTRAST (w1csex_new)=Indicator(1)
/CONTRAST (New_SN_co_ethn)=Indicator(1)
/CONTRAST (Generatie)=Indicator(1)
/CONTRAST (d3_ethnic_suri_antil)=Indicator(1)
/CONTRAST (d2_ethnic_Turks)=Indicator(1)
/CONTRAST (New_snXgeneratie)=Indicator(1)
/CONTRAST (New_snXd3_suri_antil)=Indicator(1)
/CONTRAST (New_snXd2_Turks)=Indicator(1)
/CLASSPLOT
/CASEWISE OUTLIER(2)
/PRINT=GOODFIT CI(95)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).

```

Tabel 38: Model 5 logistische regressieanalyse andere etnische referentiegroep

Variables in the Equation									
	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)		
							Lower	Upper	
Step 1 ^a									
age at time of interview	-.028	,029	,885	1	,347	,973	,918	1,030	
Dummie sekse(1)	,010	,336	,001	1	,977	1,010	,523	1,950	
Opleidingsniveau	-.145	,062	5,508	1	,019	,865	,767	,976	
Nieuwe onafhankelijke variabele			6,982	2	,030				
Nieuwe onafhankelijke variabele(1)	,993	,764	1,687	1	,194	2,698	,603	12,067	
Nieuwe onafhankelijke variabele(2)	1,864	,707	6,955	1	,008	6,451	1,614	25,781	
Generatie migranten(1)	1,185	,604	3,847	1	,050	3,269	1,001	10,678	
Wel of niet Surinaams/Antilliaans(1)	-3,340	1,182	7,983	1	,005	,035	,003	,360	
Wel of niet Turks(1)	1,302	,697	3,487	1	,062	3,676	,937	14,413	
New_snXgeneratie			,639	2	,727				
New_snXgeneratie(1)	-.043	,974	,002	1	,965	,958	,142	6,460	
New_snXgeneratie(2)	-.589	,805	,536	1	,464	,555	,115	2,687	
New_snXd3_suri_antil			1,180	2	,554				
New_snXd3_suri_antil(1)	,699	1,448	,233	1	,629	2,012	,118	34,387	
New_snXd3_suri_antil(2)	1,488	1,423	1,094	1	,296	4,427	,272	71,976	
New_snXd2_Turks			,012	2	,994				
New_snXd2_Turks(1)	17,680	4173,878	,000	1	,997	47655840,82	,000	.	
New_snXd2_Turks(2)	,095	,878	,012	1	,913	1,100	,197	6,150	
Constant	2,292	1,267	3,273	1	,070	9,895			

a. Variable(s) entered on step 1: age at time of interview, Dummie sekse, Opleidingsniveau, Nieuwe onafhankelijke variabele, Generatie migranten, Wel of niet Surinaams/Antilliaans, Wel of niet Turks, New_snXgeneratie, New_snXd3_suri_antil, New_snXd2_Turks.

Bijlage 3: Assumptietoets, outliers en multicollineariteit

In deze bijlage zullen de assumptietoetsing, outliers en multicollineariteit uitgebreid besproken worden. Om te achterhalen of de resultaten uit de analyse betrouwbaar en generaliseerbaar zijn, is het belangrijk te controleren in hoeverre de data voldoet aan de assumptie behorend bij een logistische regressieanalyse; onafhankelijke observaties. Daarnaast is het belangrijk om eventuele outliers inzichtelijk te krijgen en te controleren in hoeverre er sprake is van multicollineariteit.

Assumptietoets

Omdat de dataset verkregen is door het werk van andere onderzoekers, is het belangrijk te kijken hoe zij respondenten geselecteerd hebben. Door hiernaar te kijken kan vastgesteld worden in hoeverre de observaties onafhankelijk zijn. Omdat de NeLLS studie (Tolsma et al., 2014) een over-sample van Turkse en Marokkaanse Nederlanders bevat en mensen uit deze groepen die wonen in kleine dorpen op het platteland niet meegenomen zijn, zijn de respondenten niet geheel aselekt getrokken maar op een quasi-willekeurige wijze. Om toch onafhankelijke observaties te kunnen verkrijgen zijn 35 gemeenten op quasi-willekeurige wijze geselecteerd door de gemeenten onder te verdelen per regio en urbanisatiegraad. In de uiteindelijke steekproef moesten de vier grote westelijke steden Amsterdam, Rotterdam, Utrecht en Den Haag wel meegenomen worden om een representatieve steekproef van Turkse en Marokkaanse Nederlanders te verkrijgen. Vervolgens zijn respondenten op quasi-willekeurige wijze uit het geboorteregister getrokken op basis van etniciteit en leeftijd, weer met een over-sample van Turkse en Marokkaanse Nederlanders. Op deze manier is getracht onafhankelijke observaties te kunnen verkrijgen.

Outliers

Verder is het, voor de betrouwbaarheid en generaliseerbaarheid van de analyse, belangrijk dat er geen invloedrijke punten aanwezig zijn in de data, oftewel geen outliers. Outliers zijn punten die afwijken van de overige punten; zij kunnen veel groter of veel kleiner zijn. Een manier om outliers te detecteren is door te kijken naar de leverage waarden behorend bij alle cases uit de dataset. Deze waarden laten namelijk zien in hoeverre de cases invloed uitoefenen op de fit van de modellen. Wanneer één case veel invloed zou hebben op de

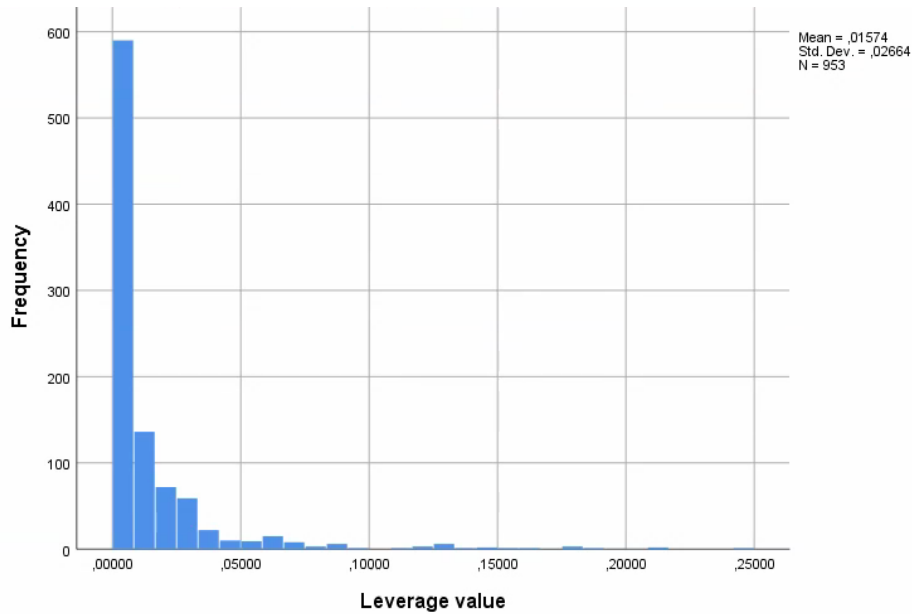
modelfit, zou deze case bepalend kunnen zijn voor de andere cases. De bovengrens van de leverage waarden kan berekend worden via de formule $3 \cdot p/n$. Hierbij staat p voor het aantal variabelen in het model en n voor de steekproefgrootte. Waarden boven deze grens worden gezien als invloedrijke punten en zijn dus mogelijk outliers. Voor het complete model 5 is dit $3 \cdot 14/953 = 0.044$.

Figuur 1 laat de frequentieverdeling zien van de leverage waarden. In totaal zijn er 69 cases die een grotere waarde dan 0.044 hebben. Deze cases zouden invloed kunnen hebben op de geschatte parameters en de fit van de modellen beïnvloed kunnen hebben. Echter, is ervoor gekozen wel alle cases op te nemen in de analyse. Leverage waarden zijn namelijk vooral relevante waarden om outliers te detecteren van continue variabelen en minder betekenisvol bij binaire logistische regressies. Dichotome variabelen zoals de afhankelijke in dit onderzoek kennen immers maar twee scores, namelijk 0 of 1. Daarbij komt dat de outliers in binaire logistische regressies juist betekenisvol kunnen zijn en in dit geval de 69 cases ongeveer gelijk scoren op het wel of niet co-etnisch getrouwd zijn als de complete steekproef (zie tabel 1).

Maar, de cases wijken wel enigszins af van de andere cases wanneer er gekeken wordt naar hun verdeling van de onafhankelijke variabele *co-etnische sociale netwerken in de wijk*. In figuur 2 is te zien dat de verdeling behorend bij de 69 cases het tegenovergestelde is van de verdeling behorend bij de complete steekproef. De meeste respondenten hebben vrijwel nooit contact met co-etnische buurtgenoten, terwijl in de complete steekproef de meeste respondenten juist wekelijks contact hebben met co-etnische buurtgenoten. Wanneer de 69 cases uit de complete steekproef gehaald zouden worden, zou het model de parameters behorend bij de onafhankelijke variabele wellicht beter kunnen schatten. Anderzijds zou de verdeling van deze variabele dan nog schever worden, aangezien er al relatief weinig respondenten zijn die vrijwel nooit of geregeld contact hebben met co-etnische buurtgenoten. In dat geval zou het lastiger worden om betrouwbare uitspraken te doen over mensen die vrijwel nooit of geregeld contact hebben met co-etnische buurtgenoten, omdat deze groepen erg klein zijn. Ook hierom is ervoor gekozen de cases wel mee te nemen in de analyse.

```
DATASET COPY BWS_Outliers.  
DATASET ACTIVATE BWS_Outliers.  
FILTER OFF.
```

```
USE ALL.
SELECT IF (LEV_2 > 0.044).
EXECUTE.
DATASET ACTIVATE DataSet1.
```



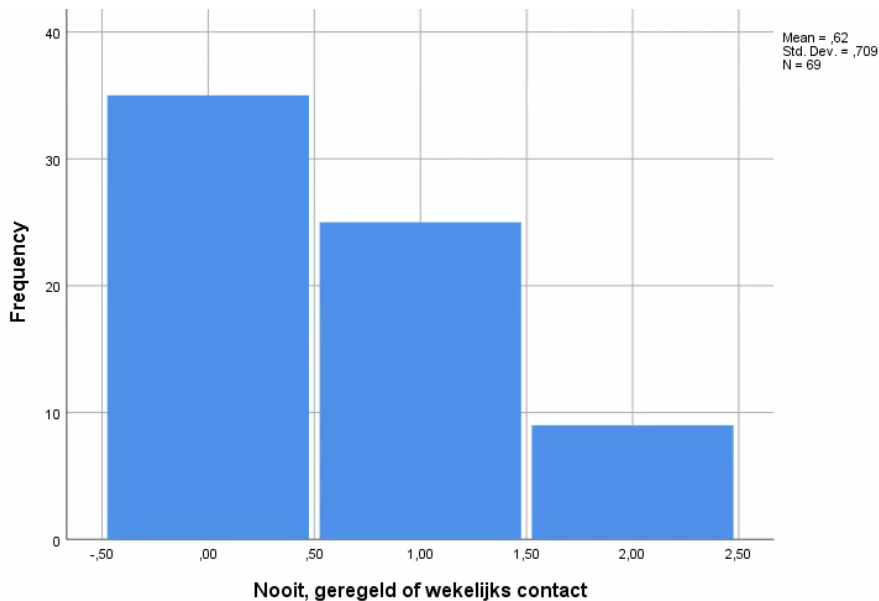
```
GRAPH
/HISTOGRAM=LEV_2.
```

Figuur 1: Frequentieverdeling van de leverage waarden

Tabel 1: SPSS-output van de frequentieverdeling wel of niet co-etnisch huwelijk

		Wel of niet co-etnisch huwelijk			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Niet van toepassing	27	39,1	39,1	39,1
	Wel van toepassing	42	60,9	60,9	100,0
Total		69	100,0	100,0	

```
FREQUENCIES
VARIABLES=New_SN_co_eth
/ORDER=ANALYSIS.
```



GRAPH
/HISTOGRAM=New_SN_co_eth.

Figuur 2: Freq. verdeling contact met co-etnische buurtgenoten 69 outliers

Multicollineariteit

Tot slot is belangrijk dat er geen sprake is van multicollineariteit, dat wil zeggen dat er geen onderlinge samenhang is tussen de variabelen in een model. Wanneer generatie en leeftijd bijvoorbeeld sterk correleren, is het lastig te achterhalen hoeveel invloed zij afzonderlijk van elkaar hebben op het al dan niet hebben van een co-etnisch huwelijk, omdat zij (gedeeltelijk) overlappen. Om te controleren in hoeverre er sprake is van multicollineariteit kan er gekeken worden naar de Variance Inflation Factor (VIF) waarden van alle variabelen in het model. In tabel 2 is een overzicht gegeven van de VIF-waarden behorend bij de variabelen uit modellen 3, 4 en 5. Wanneer een grenswaarde van 10 aangenomen wordt, zijn er enkel in model 5 een aantal variabelen die een waarde hebben die hier (ruim) boven vallen. Echter, de variabelen die samenhangen zijn de twee dummy's behorend bij *etniciteit* en de interactietermen. Omdat de twee dummy's samen één variabele vormen en de variabelen die onderdeel zijn van de interactietermen ook één variabele vormen, is het logisch dat deze VIF-waarden hoog zijn. Toch is ervoor gekozen deze variabelen op te nemen in de analyse, om zo de hypothesen te kunnen toetsen en inzichtelijk te maken hoe de verbanden kunnen lopen. Gevolgen hiervan zijn dat het schatten van de betreffende parameters bemoeilijkt kan worden en de betrouwbaarheid van deze resultaten in het geding kunnen komen. Hier zal rekening mee gehouden moeten worden tijdens het trekken van conclusies.

Tabel 2: VIF-waarden per variabele per model

Variabelen in model	Model 3	Model 4	Model 5
Leeftijd	1.230	1.239	1.240
Sekse	1.066	1.071	1.071
Opleidingsniveau	1.075	1.077	1.081
Co-etnische sociale netwerken in de wijk	1.063	4.512	28.219
Generatie	1.186	5.574	5.610
Etniciteit: Marokkaans	9.904	9.916	27.245
Etniciteit: Turks	9.932	9.934	28.264
Co-etnische sociale netwerken in de wijk X Marokkaans	-	-	49.608
Co-etnische sociale netwerken in de wijk X Turks	-	-	55.294
Co-etnische sociale netwerken in de wijk X generatie	-	9.627	9.692

Syntax:

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COLLIN TOL
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT Co_eth_huwelijk
/METHOD=ENTER w1cage w1csex_new Opleidingsniveau New_SN_co_ethn Generatie
d1_ethnic_Marokkaans
d2_ethnic_Turks.
```

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COLLIN TOL
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT Co_eth_huwelijk
/METHOD=ENTER w1cage w1csex_new Opleidingsniveau New_SN_co_ethn Generatie
d1_ethnic_Marokkaans
d2_ethnic_Turks New_snXgeneratie.
```

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COLLIN TOL
/CRITERIA=PIN(.05) POUT(.10)
```

/NOORIGIN
/DEPENDENT Co_eth_huwelijk
/METHOD=ENTER w1cage w1csex_new Opleidingsniveau New_SN_co_ethn Generatie
d1_ethnic_Marokkaans
d2_ethnic_Turks New_snXgeneratie New_snXd1_Marok New_snXd2_Turks.