

Oud geleerd, jong gedaan

Een onderzoek naar de invloed van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden

Naam: Myrthe Sinnema
Studentnummer: S3977064
Universiteit: Rijksuniversiteit Groningen
Faculteit: Gedrag- en maatschappijwetenschappen
Bachelor: Sociologie
Onderzoeksvraag: Hebben mensen wiens ouders hoogopgeleid zijn meer vertrouwen in onbekenden dan mensen wiens ouders laagopgeleid zijn?
Begeleider: Vincent Frey/Jacob Dijkstra
Datum: 07-06-2023

Inhoudsopgave

Abstract	3
Inleiding	4
Theoretisch kader	7
Een optimistische houding	7
Opgroeien in een welvarend klimaat	8
<i>Financieel vangnet</i>	8
<i>Opgroeien in een veilige buurt</i>	10
De invloed van geslacht, leeftijd en opleidingsniveau	12
<i>De invloed van geslacht</i>	12
<i>De invloed van leeftijd</i>	12
<i>De invloed van opleidingsniveau</i>	13
Methode	14
Data en procedure	14
Operationalisaties	15
<i>Vertrouwen in onbekenden</i>	15
<i>Opleidingsniveau vader en moeder</i>	15
<i>Economische situatie gezin</i>	16
<i>Geslacht en leeftijd</i>	17
<i>Opleidingsniveau respondent</i>	17
Analyse-opzet	18
Resultaten	20
Beschrijvende statistieken	20
<i>Univariate statistieken</i>	20
<i>Associatiematen</i>	22
Modevaluatie	25
<i>Onafhankelijkheid</i>	25
<i>Deviance en Hosmer en Lemeshow Goodness-of-Fit Test</i>	25
<i>Multicollineariteit</i>	26
<i>Uitbijters en invloedrijke punten</i>	26
Hypothesetoetsing	26
Conclusie	32
Discussie	35
Literatuur	38
Bijlage 1	40
Bijlage 2	50
Bijlage 3	64

Abstract

Geven hoogopgeleiden het vertrouwen dat zij hebben in onbekenden door aan hun kinderen? Het is belangrijk hier onderzoek naar te doen, omdat ouders een grote rol spelen in de vroege jeugd van hun kinderen en het is aannemelijk dat ze (een deel van) hun vertrouwen doorgeven aan hun kinderen. Opleidingsniveau lijkt sterk samen te hangen met het opleidingsniveau van iemands ouders. Zo hebben hoogopgeleide ouders vaak hoogopgeleide kinderen en laagopgeleide ouders laagopgeleide kinderen. Omdat hoogopgeleiden over het algemeen meer vertrouwen hebben in onbekenden dan laagopgeleiden (Uslaner, 2002), is het interessant om te onderzoeken of en op welke manier hoogopgeleide ouders dit vertrouwen doorgeven aan hun kinderen of dat het toch grotendeels wordt bepaald door het eigen opleidingsniveau. Om te onderzoeken of mensen met hoogopgeleide ouders meer vertrouwen hebben in onbekenden dan mensen met laagopgeleide ouders is gekeken naar een drietal verklaringen: de socialisatieverklaring, de welvaartsverklaring en de buurteffect verklaring. Om het verband tussen het opleidingsniveau van iemands ouders en het vertrouwen in onbekenden te testen is gebruik gemaakt van een bestaande dataset (EVS, 2022). Respondenten hebben vragenlijsten ontvangen en deze mondeling beantwoord of digitaal ingevuld. De data laten zien dat er geen bewijs is gevonden dat mensen met hoogopgeleide ouders meer vertrouwen hebben in onbekenden dan mensen met laagopgeleide ouders. Om die reden worden enkele implicaties besproken en enkele aanbevelingen gedaan voor vervolgonderzoek.

Inleiding

Vertrouwen is een belangrijk begrip op zowel individueel niveau als op maatschappelijk niveau. Op individueel niveau maakt vertrouwen deel uit van iemands sociaal kapitaal (Welch et al., 2005). Des te groter het sociaal kapitaal, des te meer toegang je hebt tot informatie en des te beter je om hulp kunt vragen. Dit grotere sociaal kapitaal leidt tot meer levenstevredenheid (Welch et al., 2005). Op maatschappelijk niveau zorgt vertrouwen ervoor dat een samenleving draaiende blijft. Onderling vertrouwen in een maatschappij zorgt bijvoorbeeld voor meer economische groei en meer efficiëntie op de markt (Delhey & Newton, 2003). Wanneer mensen elkaar onderling vertrouwen worden afspraken aangegaan en nageleefd, waardoor groei en efficiëntie kan worden gerealiseerd.

In de bestaande literatuur over vertrouwen worden vele concepten naast elkaar gebruikt, zoals sociaal vertrouwen, algemeen vertrouwen en horizontaal vertrouwen. Voor alle soorten vertrouwen geldt dat vertrouwen de verwachting is hoe iemand zich richting jou zal gaan gedragen (Leenheer, Gesthuizen & Savelkoul, 2021). Wanneer je verwacht dat iemand je kwaad zal doen heb je weinig vertrouwen in die persoon en wanneer je verwacht dat iemand je geen kwaad zal doen heb je veel vertrouwen in die persoon. In dit onderzoek wordt specifiek gekeken naar het vertrouwen dat mensen hebben in onbekenden. Het gaat er hierbij over in hoeverre je iemand vertrouwt die je nog nooit eerder hebt ontmoet.

Een hoog opleidingsniveau lijkt positief bij te dragen aan de mate van vertrouwen in onbekenden (Hooghe, Marien & Vroome, 2012). Zo zijn hoogopgeleiden vaak intelligenter, wat blijkt uit het behaalde opleidingsniveau waarvoor bepaalde capaciteiten benodigd zijn. Iemand die intelligent is kan de intenties en motivaties van anderen vaak beter inschatten en op die manier onbetrouwbaar gedrag signaleren (Yamagishi, Kikuchi & Kosugi, 1999). Verder zijn hoogopgeleiden vaak bevoorrecht in de maatschappij, waardoor ze beter bestand

zijn tegen onbetrouwbaar gedrag (Hooghe, 2007). Het is voor hen eenvoudiger om onbekende mensen te vertrouwen.

Hoewel meerdere onderzoeken hebben gekeken naar de invloed van opleidingsniveau op de mate van vertrouwen in onbekenden, is nog weinig bekend in hoeverre het opleidingsniveau van ouders bijdraagt aan de mate van vertrouwen van hun kinderen. Ouders vervullen een belangrijke rol in de vroege jeugd van hun kinderen en het opleidingsniveau van ouders zou daarom een grote invloed kunnen hebben. Als blijkt dat kinderen van hoogopgeleide ouders meer vertrouwen hebben in onbekenden, kan gekeken worden naar welke aspecten hierin een rol spelen en hoe laagopgeleide ouders ervoor kunnen zorgen dat hun kinderen meer vertrouwen hebben in onbekenden. Op die manier kan het vertrouwen in onbekenden onder de gehele bevolking worden gemaximaliseerd. Doordat dit onderzoek op verschillende manieren de link legt tussen de mate van vertrouwen in onbekenden van zowel ouders als hun kinderen en daarbij kijkt naar het opleidingsniveau, draagt het iets bij aan de bestaande literatuur.

Voorafgaand aan het onderzoek is het goed om kort toe te lichten wat bedoeld wordt met een laag, middelhoog en hoog opleidingsniveau. In Nederland geldt over het algemeen dat basisonderwijs, het vmbo, de eerste 3 jaren van de havo of het vwo of een mbo-1 opleiding worden beschouwd als laagopgeleid (CBS, 2019). Verder geldt dat de bovenbouw van de havo of het vwo en de niveaus mbo-1, mbo-2 en mbo-3 onder de categorie middelhoog opleidingsniveau behoren. Ten slotte valt onder een hoog opleidingsniveau een opleiding aan het hbo of aan de universiteit.

Zoals eerder benoemd wordt er in dit onderzoek gekeken naar de invloed van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden. De concrete vraag die in dit onderzoek centraal staat luidt als volgt: *Hebben mensen wiens ouders hoogopgeleid zijn meer vertrouwen in onbekenden dan mensen wiens ouders laagopgeleid*

zijn? Ik verwacht hierbij een positief verband, wat betekent dat iemand met hoogopgeleide ouders meer vertrouwen heeft in onbekenden. Dit positieve verband betekent ook dat iemand met laagopgeleide ouders juist minder vertrouwen heeft in onbekenden. Dit onderzoek richt zich grotendeels op de invloed van de economische situatie van gezinnen op de mate van vertrouwen in onbekenden. Wanneer ouders bijvoorbeeld veel inkomen hebben en hiervan gemakkelijk rond kunnen komen, neemt dit voor hun kinderen een deel van de risico's om onbekenden te vertrouwen weg.

Dit onderzoek maakt gebruik van data van de European Values Study. De vragenlijsten zijn onder andere verspreid onder de Nederlandse bevolking, waardoor de data geschikt is voor dit onderzoek. Er wordt gekeken naar het opleidingsniveau van zowel de vader als moeder, het vertrouwen in onbekenden en in hoeverre de ouders moeite hadden om rond te komen. Verder wordt er gebruik gemaakt van multiële regressie.

Theoretisch kader

Een optimistische houding

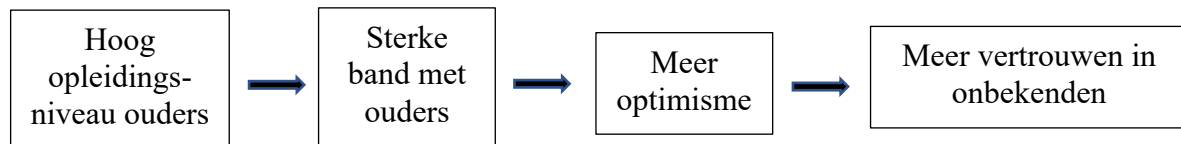
Hoogopgeleiden zijn over het algemeen werkzaam in hogere functies, hebben meer inkomen en een hogere status waardoor ze worden gezien als succesvol (Centraal Bureau voor de Statistiek, 2021). Dit leidt ertoe dat ze vaker met eerlijkheid en respect worden behandeld (Delhey & Newton, 2003). Het tegenovergestelde geldt voor laagopgeleiden, die door hun opleidingsniveau een lagere positie op de maatschappelijke ladder hebben. Ze hebben minder inkomen en een lagere status (Centraal Bureau voor de Statistiek, 2021), waardoor ze niet snel beoordeeld zullen worden als succesvol en met minder eerlijkheid en respect worden behandeld (Delhey & Newton, 2003). Zo zal in veel gevallen bijvoorbeeld een schoonmaakster van een hotel door haar lagere opleidingsniveau met minder eerlijkheid en respect worden behandeld dan de eigenaar van het hotel.

Doordat hoogopgeleiden vaak met eerlijkheid en respect worden behandeld, staan ze in een gunstige positie (Delhey & Newton, 2003) en hebben ze vaak een optimistischere houding richting anderen (O'Connor & Graham, 2019). Ze schatten situaties eerder positief dan negatief in. Het is aannemelijk dat hoogopgeleiden door deze optimistische kijk op de wereld eerder geneigd zijn om onbekenden te vertrouwen. Ze verwachten namelijk dat een onbekende persoon eerder goede dan slechte intenties heeft.

Ouders spelen een belangrijke rol in de vroege jeugd van hun kinderen (Delhey & Newton, 2003) en maken deel uit van hun informele netwerk. Dit informele netwerk speelt een grote rol in het aanleren van vertrouwen in anderen (Leenheer et al., 2021).

Hoogopgeleiden hebben vaak een sterkere band met hun kinderen dan laagopgeleiden waardoor kinderen van hoogopgeleiden eerder geneigd zijn om hun ouders als rolmodel te zien (Farrington, 2017). Ze nemen hierdoor de optimistische kijk van hun ouders over en zijn zelf ook sneller geneigd om onbekenden te vertrouwen. Wanneer hoogopgeleide ouders dus

een optimistische kijk hebben, dan is de kans aannemelijk dat hun kinderen dit optimisme overnemen en daardoor in hun latere leven meer vertrouwen hebben in onbekenden. Deze verklaring wordt de socialisatieverklaring genoemd en is schematisch weergegeven in figuur 1.



Figuur 1: grafische weergave van de socialisatieverklaring

De hypothese die volgt uit de socialisatieverklaring luidt als volgt: *Naarmate men ouders heeft die hoogopgeleid zijn heeft men meer vertrouwen in onbekenden.*

Opgroeien in een welvarend klimaat

Financieel vangnet

In Nederland verdienen hoogopgeleide Nederlanders gemiddeld bijna twee keer zoveel als laagopgeleide Nederlanders (Centraal Bureau voor de Statistiek, 2011; Zwart, 2021).

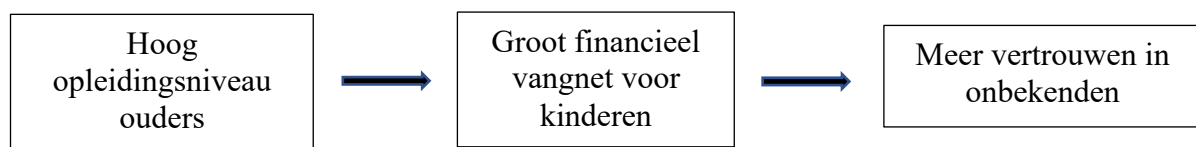
Hoogopgeleide Nederlanders komen bijvoorbeeld vaker terecht in leidinggevende functies, die gepaard gaan met een hoger inkomen. Dit hogere inkomen zou ertoe kunnen leiden dat hoogopgeleiden minder moeite hebben om rond te komen.

Uit onderzoek blijkt dat mensen met een hoger inkomen meer vertrouwen hebben in onbekenden (Leenheer et al., 2021). Een verklaring hiervoor is dat door een groter financieel vangnet minder risico's verbonden zijn aan het vertrouwen van onbekende mensen (Delhey & Newton, 2003). Omdat bij gezinnen waarbij één of beide ouders hoogopgeleid zijn het huishoudinkomen hoger is en kinderen hiervan meeprofiteren, geldt dit financiële vangnet voor zowel de ouders als de kinderen. Kinderen zijn door dit financiële vangnet minder bang

om risico's te nemen, omdat ze weten dat ze terug kunnen vallen op het financieel kapitaal van hun ouders.

Wanneer je als kind op de middelbare school tijdens een studiesessie bijvoorbeeld naar de kantine gaat maar je laptop niet mee wilt nemen, ben je eerder geneigd een onbekende te vertrouwen om even op te letten wanneer je een financieel vangnet hebt. Je kunt het risico dat je laptop wordt gestolen namelijk veroorloven, omdat je terug kunt vallen op het financieel kapitaal van je ouders. Mocht je dit vangnet niet hebben omdat je ouders door een lager inkomen relatief weinig financieel kapitaal hebben, dan kun je je het risico niet veroorloven. Je bent in zo'n geval eerder geneigd een onbekende te wantrouwen en je laptop met je mee te nemen. Het risico dat je laptop gestolen wordt is namelijk te groot.

Mochten iemands ouders dus een relatief hoog inkomen hebben gehad in diegene zijn of haar jeugd, dan is de kans groot dat die persoon meer geneigd was onbekenden te vertrouwen. Omdat dit vertrouwen wordt opgebouwd en aangeleerd in de vroege jeugd (Delhey & Newton, 2003) is het aannemelijk dat iemand op latere leeftijd ook meer vertrouwen heeft in onbekenden. Op die manier zorgt een hoog opleidingsniveau van ouders er via het inkomen van deze ouders voor dat er een groter financieel vangnet is voor de kinderen welke de risico's van het vertrouwen van onbekenden opvangt. Deze verklaring wordt de welvaartsverklaring genoemd en is schematisch weergegeven in figuur 2.



Figuur 2: grafische weergave van de welvaartsverklaring

Opgroeien in een veilige buurt

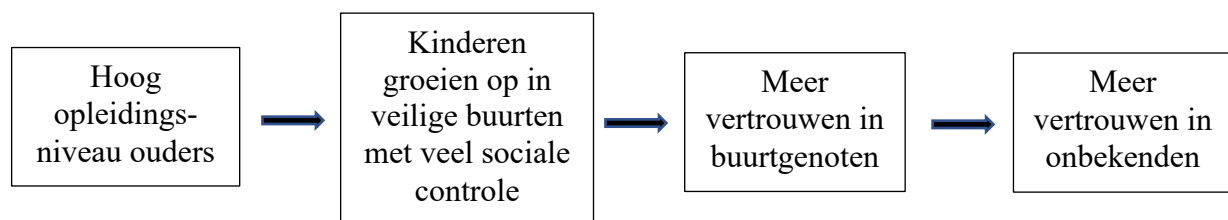
Zoals eerder aangegeven hebben hoogopgeleiden vaak een hoger inkomen dan laagopgeleiden. Gezinnen waarbij één of beide ouders hoogopgeleid zijn gaan daardoor vaker in buurten wonen waar de huizen relatief meer kosten (Ham, Boschman & Vogel, 2018). Andersom geldt hetzelfde: gezinnen waarbij één of beide ouders laagopgeleid zijn gaan vaker in buurten wonen waar de huizen relatief weinig kosten, omdat hun inkomen de bekostiging van de duurdere huizen niet toelaat.

Onderzoek wijst uit dat er meer criminaliteit voorkomt in armere buurten dan in rijkere buurten (Galster, 2002). Wanneer er enige vorm van criminaliteit plaatsvindt in een buurt, is de kans aannemelijk dat de criminele cultuur in die buurt dominant wordt (Galster, 2002). Dit heeft te maken met een drietal factoren: het aandeel niet-criminele activiteiten neemt af, minder mensen uit de buurt zullen de delicten rapporteren bij de politie en criminaliteit in de buurt zal eerder worden genormaliseerd. Door deze factoren is er sprake van een neerwaarts spiraal waarbij er in buurten waar al enige vorm van criminaliteit is, het aandeel criminaliteit steeds groter wordt.

Andersom geldt dat er relatief minder criminele activiteiten plaatsvinden in rijkere buurten. In deze buurten is veel sociale controle, waarbij buurtbewoners de veiligheid van jongeren reguleren, lastige situaties gezamenlijk proberen op te lossen en laten zien vertrouwen in elkaar te hebben (Farrington, 2017). Op die manier wordt het gedrag van de jongeren in de gaten gehouden en worden er indien nodig maatregelen getroffen. Deze sociale controle draagt eraan bij dat criminele delicten onder de jeugd geen onderdeel worden van de dominante cultuur in de buurt.

Wanneer jongeren opgroeien in buurten waar relatief weinig criminaliteit voorkomt, hebben zij meer vertrouwen in buurtgenoten (Farrington, 2017). Opgroeien in een veilige buurt zorgt voor goede ervaringen in het vertrouwen van anderen en dus in een toename van

vertrouwen in buurtgenoten. Hoewel dit vertrouwen specifiek gaat over vertrouwen in buurtgenoten, kan door middel van het *spill-over effect* worden verklaard waarom deze jongeren meer vertrouwen in onbekenden hebben. Volgens het spill-over effect zijn mensen eerder geneigd om onbekenden te vertrouwen wanneer ze veel vertrouwen hebben in specifieke mensen (Leenheer et al., 2021), in dit geval buurtgenoten. Omdat de basis voor de mate van vertrouwen wordt gelegd in de vroege jeugd en niet veel meer verandert in de latere jaren (Delhey & Newton, 2003), kan ervan uit worden gegaan dat de jongeren uit rijkere buurten op latere leeftijd nog steeds meer vertrouwen hebben in onbekenden. Zo zorgt een hoog opleidingsniveau van ouders er dus via de economische situatie van het gezin voor dat kinderen opgroeien in een veilige buurt met veel sociale controle, waardoor de jongeren meer vertrouwen hebben in buurtgenoten. Door middel van het spill-over effect hebben deze jongeren meer vertrouwen hebben in onbekenden. Deze verklaring wordt de buurteffect verklaring genoemd en is schematisch weergegeven in figuur 3.



Figuur 3: grafische weergave van de buurteffect verklaring

Voordat uitspraken kunnen worden gedaan over eventuele voordelen van de economische situatie van gezinnen met een of meerdere hoogopgeleide ouders op de mate van vertrouwen in onbekenden, moet vooraf kunnen worden vastgesteld dat een hoog opleidingsniveau daadwerkelijk leidt tot een betere economische situatie. Om deze reden is de volgende hypothese tot stand gekomen: *Gezinnen waarbij de ouders hoogopgeleid zijn hadden vroeger minder moeite om rond te komen dan gezinnen waarbij de ouders laagopgeleid zijn.*

Aan de hand van de welvaartsverklaring en de buurteffect verklaring wordt vervolgens gekeken of de economische situatie van een gezin invloed heeft op het verband tussen het opleidingsniveau van iemands ouders en de mate van vertrouwen in onbekenden. Hierbij hoort de volgende hypothese: *Het positieve effect van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden wordt gedeeltelijk verklaard door de economische situatie van het gezin.*

De invloed van geslacht, leeftijd en opleidingsniveau

De invloed van geslacht

Het is aannemelijk dat geslacht invloed heeft op de mate van vertrouwen in onbekenden. Aan de ene kant wijst onderzoek uit dat vrouwen minder vertrouwen hebben in onbekenden omdat ze meer risico lopen om te worden gediscrimineerd (Woolcock, 1998; Delhey & Newton, 2003). Aan de andere kant zijn er ook onderzoeken die het tegenovergestelde laten zien: vrouwen blijkt juist méér vertrouwen hebben in onbekenden dan mannen (Feingold, 1994). Er lijken dus verschillen te zijn tussen mannen en vrouwen wat betreft de mate van vertrouwen in onbekenden, maar toch is dit verschil niet erg groot in westerse landen (Rontos & Roumeliotou, 2013). Hoewel geslacht geen invloed heeft op het opleidingsniveau van ouders, is het erg belangrijk om geslacht als controlevariabele op te nemen in het model. Geslacht heeft een dusdanig grote invloed op de mate van vertrouwen in onbekenden, dat er betere schattingen kunnen worden gedaan wanneer geslacht wordt opgenomen in het model.

De invloed van leeftijd

De gemiddelde leeftijd van Nederlanders die mee hebben gedaan aan het onderzoek is drieënvijftig jaar (EVS, 2022). Aan de hand van bestaande literatuur kan er kan geen eenduidig antwoord worden gegeven op de vraag of ouderen juist meer of minder vertrouwen

hebben in onbekenden dan jongeren. Aan de ene kant lijkt het erop dat ouderen meer vertrouwen hebben in onbekenden dan jongeren (Uslaner, 1998). Dit zou er mee te maken kunnen hebben dat ze zijn opgegroeid in een veiligere omgeving dan eerdere generaties. Aan de andere kant is er ook bewijs gevonden voor de zogenoemde *u-curve*. Dit houdt in dat zowel jongeren als ouderen relatief weinig vertrouwen hebben in onbekenden en dat de groep hier tussenin relatief veel vertrouwen heeft in onbekenden (Rontos & Roumeliotou, 2013). Hoewel leeftijd – net als geslacht – geen invloed heeft op het opleidingsniveau van ouders, is het belangrijk om leeftijd als controlevariabele op te nemen in het model. Op die manier wordt voorkomen dat een schijnrelatie tussen het opleidingsniveau van ouders en de mate van vertrouwen in onbekenden wordt gemeten.

De invloed van opleidingsniveau

In het onderzoek wordt gekeken naar de invloed van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden. Omdat iemands opleidingsniveau sterkt lijkt samen te hangen met zowel het opleidingsniveau van de ouders (hoogopgeleide ouders hebben vaak hoogopgeleide kinderen; laagopgeleide ouders hebben vaak laagopgeleide kinderen) als met de mate van vertrouwen in onbekenden, is het belangrijk om opleidingsniveau op te nemen als controlevariabele in het model. We kunnen op die manier onderzoeken of het opleidingsniveau van iemands ouders een effect heeft op vertrouwen, boven op het effect dat het eigen opleidingsniveau heeft. Uit onderzoek blijkt namelijk dat hoogopgeleiden over het algemeen een hogere mate van vertrouwen hebben in onbekenden dan laagopgeleiden (Uslaner, 2002). Door opleidingsniveau op te nemen in het model wordt voorkomen dat een schijnrelatie tussen het opleidingsniveau van ouders en de mate van vertrouwen in onbekenden wordt gemeten.

Methode

Data en procedure

Dit onderzoek maakt gebruik van de European Values Study (EVS). Het doel van dit programma is om meer inzicht te krijgen in onderwerpen als familie, werk, religie, politiek en maatschappij specifiek voor inwoners in Europa (EVS, 2022). EVS is een onderzoeksprogramma dat inmiddels wordt uitgevoerd in 37 landen in Europa, waaronder Nederland. Het onderzoek is opgezet in 1981 en is vanaf dat jaar om de negen jaar herhaald. Omdat 2017 het meest recente jaar is waarin onderzoek is gedaan, is ervoor gekozen om slechts te kijken naar verzamelde data uit dat jaar. Het onderzoek over de invloed van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden richt zich op volwassen Nederlanders. Er is daarom een selectie gemaakt voor Nederlanders in de steekproef.

De data zijn verzameld op basis van een gestratificeerde steekproef. Dit houdt in dat de populatie is opgedeeld in subpopulaties, in dit geval landen. Vervolgens zijn binnen deze landen aselecte steekproeven uitgevoerd. Er is gebruik gemaakt van twee verschillende meetmethoden: *Computer Assisted Personal Interviewing* (CAPI) en *Computer Assisted Web Interviewing* (CAWI). CAPI houdt in dat er een face-to-face interview is afgenomen waarbij assistentie door een computer wordt verleend. Wanneer respondenten de vraag bijvoorbeeld niet begrijpen, kan de interviewer verduidelijking geven. CAWI houdt in dat er een vragenlijst naar de respondent wordt gestuurd en de respondent kan deze vervolgens invullen en terugsturen. Bij de CAWI-methode is gebruik gemaakt van een matrix design. Dit houdt in dat respondenten slechts een blok met vragen krijgen in plaats van de gehele vragenlijst. Wanneer ze de vragenlijst hebben ingevuld volgt er een tweede vragenlijst die kan worden gezien als de follow-up vragenlijst.

In totaal zijn 2739 volwassen Nederlanders benaderd om deel te nemen aan het onderzoek. Hiervan zijn 686 respondenten face-to-face geïnterviewd (CAPI) en hebben 2053 respondenten de vragenlijst per mail ontvangen (CAWI). Uiteindelijk hebben van deze 2053 respondenten 324 respondenten de follow-up vragenlijst niet ingevuld, waardoor van hen geen complete data beschikbaar is. Verder hebben 11 respondenten minder dan 50% van de vragen in de follow-up vragenlijst ingevuld, waardoor ook zij niet zijn meegenomen in de data. In totaal is er data beschikbaar van 2404 Nederlanders uit het jaar 2017.

Operationalisaties

Vertrouwen in onbekenden

De afhankelijke variabele in dit onderzoek is *vertrouwen in onbekenden* en de vraag die hierbij is gesteld luidt als volgt: ‘In welke mate vertrouwt u mensen die u voor het eerst ontmoet?’. Respondenten konden kiezen uit een 4-puntsschaal waarbij 1 = vertrouw ik volledig en 4 = vertrouw ik helemaal niet.

Veruit de meeste mensen (69.1%) hebben geantwoord dat ze over het algemeen een beetje vertrouwen hebben in onbekenden. Omdat slechts 183 respondenten (4.1%) ‘vertrouw ik helemaal niet’ hebben ingevuld en 97 respondenten (7.8%) ‘vertrouw ik helemaal’ is ervoor gekozen om van de variabele *vertrouwen* een dichotome variabele te maken. Hierbij geldt dat 0 = geen vertrouwen in onbekenden en 1 = wel vertrouwen in onbekenden. De scores -10 tot en met -1 zijn gehercodeerd als system missing.

Opleidingsniveau vader en moeder

De onafhankelijke variabelen in dit onderzoek zijn het opleidingsniveau van vaders en het opleidingsniveau van moeders. De vragen die zijn gesteld luiden als volgt: ‘Wat is het hoogst voltooide opleidingsniveau van uw vader?’ en ‘Wat is het hoogst voltooide opleidingsniveau

van uw moeder?'. Hierbij konden de respondenten kiezen uit 18 antwoordcategorieën die variëren tussen 'lagere school niet voltooid' en 'doctoraat/gepromoveerd'. In dit onderzoek is gebruikgemaakt van gehercodeerde variabelen die door de onderzoekers al direct in het databestand zijn gezet. De categorieën van deze gehercodeerde variabelen zijn lager-, middelbaar- en hoger onderwijs. Van de antwoordcategorieën zijn dummyvariabelen gemaakt zodat ze in het model meegenomen kunnen worden. Er is gekozen om een laag opleidingsniveau als referentie te nemen, waardoor een middelhoog opleidingsniveau en hoog opleidingsniveau in het model worden opgenomen. Verder zijn de scores -10 tot en met -1 zijn gehercodeerd als system missing.

Economische situatie gezin

De variabele die hoort bij de economische situatie van het gezin gaat over de vraag of de ouders moeite hadden te eindjes aan elkaar te knopen. De vraag die is gesteld luidt als volgt: 'Geeft de beschrijving goed aan hoe uw ouders waren toen u 14 was? Mijn ouder(s) had(den) moeite de eindjes aan elkaar te knopen.' Hierbij geldt dat 1 = ja, 2 = enigszins, 3 = een klein beetje en 4 = nee. De variabele wordt vanaf nu moeite met rondkomen genoemd. Van de variabele zijn dummyvariabelen gemaakt zodat de variabele kan worden meegenomen in het model. Er is gekozen om moeite met rondkomen als referentie te nemen, waardoor een beetje moeite met rondkomen, enigszins moeite met rondkomen en geen moeite met rondkomen in het model worden opgenomen. Verder zijn de scores -10 tot en met -1 gehercodeerd als system missing.

Geslacht en leeftijd

De eerste controlevariabele in het onderzoek over vertrouwen in onbekenden is *geslacht*. De variabele is gehercodeerd waarbij geldt dat 0 = man en 1 = vrouw. De scores van respondenten die niet binnen deze categorieën vallen zijn gehercodeerd tot missende waarden. De tweede controlevariabele is *leeftijd*, welke gemeten is in jaren. De variabele loopt van 18 tot 82, waarbij 82 staat voor 82 jaar en ouder. Alle respondenten hebben hun leeftijd opgegeven waardoor er geen missende data is. Er hoeft om die reden geen hercodering te worden uitgevoerd.

Opleidingsniveau respondent

De derde controlevariabele is *opleidingsniveau respondent* en de vraag die aan respondenten is gesteld luidt als volgt: ‘Wat is het hoogste opleidingsniveau dat u voltooid heeft?’. Ook hier (net als bij het opleidingsniveau van de ouders) konden de respondenten kiezen uit 18 antwoordcategorieën die variëren tussen ‘lagere school niet voltooid’ en ‘doctoraat/gepromoveerd’. In dit onderzoek is gebruikgemaakt van een gehercodeerde variabele die ook hier al is opgezet door de onderzoekers zelf. Het opleidingsniveau van de respondenten is gecategoriseerd op een 3-puntsschaal: 1 = lager onderwijs, 2 = middelbaar onderwijs en 3 = hoger onderwijs. Van de antwoordcategorieën zijn dummyvariabelen gemaakt zodat ze in het model meegenomen kunnen worden. Er is gekozen om een laag opleidingsniveau als referentie te nemen, waardoor een middelhoog opleidingsniveau en hoog opleidingsniveau in het model worden opgenomen. Verder zijn de scores -10 tot en met -1 zijn gehercodeerd als system missing.

Analyse-opzet

In dit onderzoek worden de hypothesen getoetst door middel van logistische regressie omdat de afhankelijke variabele vertrouwen in onbekenden een dichotome variabele is. In het eerste model worden de variabelen vertrouwen in onbekenden als afhankelijke variabele en de controlevariabelen toegevoegd. Op die manier kan worden gekeken naar de invloed van de controlevariabelen op de afhankelijke variabele.

In het tweede model worden opnieuw de variabelen vertrouwen in onbekenden als afhankelijke variabele en de controlevariabelen toegevoegd, maar dit keer met het opleidingsniveau van de vader en het opleidingsniveau van de moeder als onafhankelijke variabelen. Aan de hand van dit model wordt de eerste hypothese getoetst die stelt dat iemand meer vertrouwen heeft in onbekenden wanneer deze persoon hoogopgeleide ouders heeft.

In het derde model worden het opleidingsniveau van de vader en het opleidingsniveau van de moeder als onafhankelijke variabelen toegevoegd en de economische situatie van het gezin als afhankelijke variabele. Ook hier worden de controlevariabelen aan het model toegevoegd. Het derde model wordt geschat aan de hand van lineaire regressie, omdat de variabele economische situatie bestaat uit vier categorieën.

Vervolgens wordt in het vierde model de economische situatie van het gezin toegevoegd als mediërende variabele aan het eerste model. Zowel het derde als vierde model worden gebruikt om de tweede hypothese te toetsen die stelt dat het positieve effect van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden gedeeltelijk wordt verklaard door het inkomen van de ouders.

De modelfit van het logistische regressiemodel zal allereerst worden bekeken aan de hand van de chi-kwadraattoets. Deze toets vergelijkt het model met alleen de constante en het model waaraan de variabelen zijn toegevoegd. Verder zal de modelfit beoordeeld worden

door te kijken naar de Hosmer en Lemeshow Goodness-of-Fit Test. Dit is een toets die gebruikt wordt om te beoordelen of de gebruikte data goed bij het model past.

De assumpties lineariteit, homoscedasticiteit en normaliteit worden automatisch geschonden bij logistische regressie, omdat de afhankelijke variabele een dichotome variabele is. Wel wordt er gekeken naar onafhankelijkheid als belangrijke assumptie voor het logistische regressiemodel. Hierbij wordt gekeken of de steekproef aselekt is en of er sprake is van onafhankelijke waarnemingen.

Ten slotte wordt de model diagnostiek uitgevoerd. Hierbij wordt gekeken of er sprake is van multicollineariteit, wat betekent dat er (te grote) samenhang is tussen de variabelen. De multicollineariteit wordt beoordeeld door VIF-scores uit te rekenen. Verder wordt gekeken naar invloedrijke punten die de uitkomsten van het model kunnen beïnvloeden. Dit wordt gedaan aan door leverage scores en DFBETA-scores uit te rekenen.

Resultaten

Beschrijvende statistieken

Univariate statistieken

Tabel 1 op de volgende pagina laat de beschrijvende statistieken zien van alle variabelen in het onderzoek. De uitwerking is te vinden in bijlage 1. Wat direct opvalt is dat de meeste respondenten (78.4%) hebben aangegeven vertrouwen te in onbekenden, terwijl een kleiner deel (21.6%) over het algemeen geen vertrouwen heeft in onbekenden. Voor de analyse is dit geen probleem, omdat er data beschikbaar is over 1968 respondenten en er een groep overblijft van 425 respondenten die over het algemeen geen vertrouwen hebben in onbekenden.

Verder hebben 420 respondenten (21.3%) aangegeven dat hun vaders hoogopgeleid zijn en 262 respondenten (13.3%) dat hun moeders hoogopgeleid zijn. Dit is een redelijk verschil met het percentage laagopgeleide vaders (64.3%) en laagopgeleide moeders (74.0%). Hoewel het merendeel van de respondenten laagopgeleide ouders heeft, vormt dit geen probleem vanwege de redelijk grote steekproef.

Daarnaast hebben de meeste respondenten (49.6%) aangegeven dat het gezin geen moeite had met rondkomen. De overige respondenten vormen samen de categorie waarbij het gezin tenminste enige moeite had met rondkomen en deze groep bestaat uit 992 respondenten (50.4%).

Het geslacht van de respondenten is redelijk gelijk verdeeld, maar met iets meer vrouwen (52.8%) in de steekproef dan mannen (47.2%). Verder is de leeftijd van de respondenten niet geheel normaal verdeeld ($M=53.02$; $SD=17.32$), maar dit heeft er waarschijnlijk mee te maken dat respondenten ouder dan 82 jaar wordt genoteerd als 82 jaar. Ten slotte is het opleidingsniveau van de respondenten iets gelijkverdeelddan het opleidingsniveau van hun ouders, met als grootste groep de hoogopgeleiden (45.2%).

Tabel 1: beschrijving van de opgenomen variabelen in de analyse: gemiddelde, standaarddeviatie, minimum, maximum en aantal respondenten

Variabele	Gemiddelde (standaarddeviatie) ^a	Minimum	Maximum	N totaal
Vertrouwen in onbekenden (0=geen vertrouwen; 1=wel vertrouwen)	21.6% (geen) 78.4% (wel)	0	1	1968
Opleidingsniveau vader (1=laag; 2=middelhoog; 3=hoog)	64.3% (laag) 14.4% (middelhoog) 21.3% (hoog)	1	3	1968
Opleidingsniveau moeder (1=laag; 2=middelhoog; 3=hoog)	74% (laag) 12.7% (middelhoog) 13.3% (hoog)	1	3	1968
Economische situatie gezin (1=moeite rondkomen; 2=enigszins moeite rondkomen; 3=beetje moeite rondkomen; 4=geen moeite rondkomen)	12.7% (moeite) 19.8% (enigszins) 18.0% (beetje) 49.6% (geen moeite)	1	4	1968
Geslacht (0=man; 1=vrouw)	47.2% (man) 52.8% (vrouw)	0	1	1968
Leeftijd	53.20 (17.32)	18	82	1968
Opleidingsniveau respondent (1=laag; 2=middelhoog; 3=hoog)	27.7% (laag) 27.1% (middelhoog) 45.2% (hoog)	1	3	1968

^a Bij nominale variabelen is de frequentieverdeling weergegeven in percentages

Associatiematen

Tabel 2 op pagina 24 geeft de correlaties tussen de verschillende variabelen weer. Deze correlaties zijn uitgerekend door Cramer's V scores uit te rekenen en kruistabellen uit te draaien, Pearson correlaties te berekenen en door gebruik te maken van de ANOVA-tabel. In bijlage 2 is te zien hoe de correlaties zijn uitgerekend.

De grootste samenhang is gevonden tussen het opleidingsniveau van vaders en moeders ($r=0.465$; $p<0.001$). Tabel 2.9 in bijlage 2 laat zien dat het gaat om een significant positief verband en dat de samenhang met name geldt voor laagopgeleide vaders en moeders. Het grootste deel van de respondenten (59.7%) heeft namelijk zowel een laagopgeleide vader als moeder. Een kleiner deel van de respondenten (9.9%) heeft zowel een hoogopgeleide vader als moeder.

Daarnaast heeft leeftijd een significant effect op de andere variabelen. Er is onder andere een positieve samenhang gevonden tussen het opleidingsniveau van respondenten en de mate van vertrouwen in onbekenden ($r=0.144$; $p<0.001$). Dit betekent dat respondenten met een hoog opleidingsniveau over het algemeen meer vertrouwen hebben in onbekenden dan respondenten met een laag opleidingsniveau.

Verder is een kleine samenhang gevonden tussen het opleidingsniveau van respondenten en het opleidingsniveau van vaders ($r=0.289$; $p<0.001$) en tussen het opleidingsniveau van respondenten en het opleidingsniveau van moeders ($r=0.229$; $p<0.001$). Dit is een positief verband, wat betekent dat hoogopgeleide respondenten zelf vaak hoogopgeleide ouders hebben. In totaal heeft 35.2% van de hoogopgeleide respondenten een hoogopgeleide vader en 20.9% een hoogopgeleide moeder.

Ook is er enige samenhang tussen de economische situatie van het gezin en het opleidingsniveau van vaders ($r=0.225$; $p<0.001$) en tussen de economische situatie van het gezin en het opleidingsniveau van moeders ($r=0.173$; $p<0.001$). Dit is een positief verband,

wat betekent dat hoogopgeleide vaders en moeders minder vaak moeite hadden met rondkomen vroeger. Van de respondenten met een hoogopgeleide vader had 70,7% geen moeite met rondkomen vroeger en van de respondenten met een hoogopgeleide moeder had 67.6% geen moeite met rondkomen vroeger.

Ten slotte is een klein maar significant verband gevonden tussen het opleidingsniveau van respondenten en de mate van vertrouwen in onbekenden ($r=0.092$; $p<0.001$). Van alle respondenten die hebben aangegeven vertrouwen te hebben in onbekenden, is het merendeel (47.8%) hoogopgeleid.

Tabel 2: product-momentcorrelaties van alle variabelen opgenomen in de analyse

	1. Vertr. onbekenden	2. Opleiding vader	3. Opleiding moeder	4. Econ. situatie gezin	5. Geslacht	6. Leeftijd	7. Opleiding respondent
1. Vertrouwen in onbekenden	-						
2. Opleidingsniveau vader	0.028	-					
3. Opleidingsniveau moeder	0.005	0.465**	-				
4. Economische situatie gezin	0.049	0.225**	0.173**	-			
5. Geslacht (0=man; 1=vrouw)	0.021	0.024	0.021	0.022	-		
6. Leeftijd	0.144**	0.307**	0.413**	0.210**	-0.047*	-	
7. Opleidingsniveau respondent	0.092**	0.289**	0.229**	0.150**	0.045*	0.327**	-

Modevaluatie

Onafhankelijkheid

Onafhankelijkheid is een belangrijke assumptie die gecontroleerd moet worden in een regressieanalyse. Het gaat er hierbij om of de steekproef aselekt is toegewezen en of er sprake is van onafhankelijke waarnemingen. Met een aselechte steekproef bedoelen we dat respondenten willekeurig gekozen zijn. In het onderzoek van EVS is dit het geval, omdat willekeurige Nederlanders zijn gevraagd deel te nemen aan het onderzoek. Ook is er in het onderzoek van EVS sprake van onafhankelijke waarnemingen, wat betekent dat er geen samenhang is tussen de respondenten. Er is bijvoorbeeld geen getrapte steekproef geweest en er zijn niet meerdere metingen gedaan bij dezelfde respondent.

Deviance en Hosmer en Lemeshow Goodness-of-Fit Test

Voor de modevaluatie is gekeken naar de deviance en naar de Hosmer en Lemeshow Goodness-of-Fit Test. De deviance is bekeken door chi-kwadratescores uit te rekenen voor de verschillende modellen en te kijken naar de vrijheidsgraden. Als de chi-kwadratescore lager wordt naar mate een parameter wordt toegevoegd aan het model, betekent dit dat het ‘completere’ model de afhankelijke variabele beter kan schatten. In tabel 3.1, 3.2 en 3.3 in bijlage 3 is te zien dat model 4 een lagere chi-kwadratescore heeft dan model 1 en 2. Dit betekent dat we het vertrouwen in onbekenden beter kunnen schatten wanneer de economische situatie van het gezin als variabele aan het model wordt toegevoegd.

De Hosmer en Lemeshow Goodness-of-Fit Test is een toets die wordt gebruikt om te kijken of het logistische model goed bij de data past. De nulhypothese die bij deze toets hoort is dat er geen verschil is tussen de geobserveerde en voorspelde scores. Wanneer de nulhypothese niet wordt verworpen, betekent dit dat het model goed bij de data past. Zowel model 1, model 2 als model 3 heeft een significantieniveau >0.05 . We kunnen de

nulhypothese daarom niet verwerpen en nemen aan dat de modellen goed bij de data passen.

De uitwerking is te zien in bijlage 3.

Multicollineariteit

Het is belangrijk om te kijken naar multicollineariteit, omdat dit iets zegt over de onderlinge samenhang tussen variabelen. Wanneer er een grote samenhang is kan dit ervoor zorgen dat het lastig te beoordelen is wat een variabele precies bijdraagt aan het voorspellen van de afhankelijke variabele. Daarnaast wordt de kwaliteit van het model vaak niet per se beter met het toevoegen van de variabele met grote samenhang, omdat er sprake is van onderlinge concurrentie. De variabelen verklaren dus grotendeels hetzelfde deel. De multicollineariteit is berekend aan de hand van VIF-scores, die te zien zijn in tabel 3.5 in bijlage 3. Aan de hand van de berekende scores concluderen we dat er geen grote samenhang is tussen de variabelen en er lijkt dus geen sprake van multicollineariteit.

Uitbijters en invloedrijke punten

Om te kijken of er invloedrijke punten zijn scores berekend voor leverage en DFBETA. De uitwerking staat weergegeven in bijlage 3. Er zijn geen invloedrijke punten gevonden voor de DFBETA van de constante, dus deze zijn niet in een tabel weergegeven. Ook zijn er geen invloedrijke punten gevonden voor de leverage scores.

Hypothesetoetsing

Om iets te zeggen over de hypothesen zijn verschillende regressieanalyses uitgevoerd. De uitkomsten hiervan staan weergegeven in tabel 3 op de volgende pagina.

Tabel 3: Resultaten van een regressieanalyse met enerzijds vertrouwen in onbekenden en anderzijds rondkomen als afhankelijke, opleidingsniveau vader en opleidingsniveau moeder als onafhankelijke en rondkomen als mediërende variabele

	Model 1 ^a			Model 2 ^a			Model 3 ^b		Model 4 ^a		
	b	SE	OR	b	SE	OR	b	SE	b	SE	OR
Intercept	-0.54*	0.24	0.58	-0.90*	0.31	0.41	2.98**	0.13	-0.81*	0.32	0.44
Geslacht (0=man)	0.16	0.11	1.17	0.16	0.11	1.17	0.02	0.05	0.16	0.11	1.17
Leeftijd	0.03**	0.00	1.03	0.03**	0.00	1.03	0.01**	0.00	0.03**	0.00	1.03
Middelhoog opgeleid respondent	0.36*	0.15	1.44	0.32**	0.15	1.37	0.20	0.07	0.30**	0.15	1.35
Hoogopgeleid respondent	0.78**	0.14	2.18	0.67**	0.15	1.99	0.22**	0.06	0.66	0.15	1.94
Middelhoog opgeleid vader				0.17	0.19	1.19	0.46**	0.07	0.13	0.19	1.13
Hoogopgeleid vader				0.09	0.19	1.10	0.55**	0.07	0.02	0.19	1.02
Middelhoog opgeleid moeder				0.18	0.20	1.19	0.15	0.08	0.17	0.20	1.18
Hoogopgeleid moeder				0.26	0.21	1.30	-0.00	0.09	0.27	0.22	1.31
Moeite rondkomen									-0.36*	0.18	0.70
Enigszins moeite rondkomen									-0.25	0.16	0.78
Beetje moeite rondkomen									-0.02	0.16	0.98
Deviance	1980.14			1975.46					1969.61		
R ²							0.11				
F-toets							30.54	p = 0.000			
N	1968			1968			1968		1968		

*significant bij p<0.05. ** significant bij p<0.01

^a afhankelijke variabele is vertrouwen in onbekenden, ^b afhankelijke variabele is economische situatie gezin

Om de invloed van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden te onderzoeken, zijn vier analyses uitgevoerd. De eerste analyse die is uitgevoerd is een logistische regressieanalyse met daarin de controlevariabelen geslacht, leeftijd en het opleidingsniveau van de respondent. Vertrouwen in onbekenden is hier de afhankelijke variabele. Het effect van geslacht is positief ($b=0.166$; $p=0.163$) en de bijbehorende odds ratio is 1.17. Inhoudelijk betekent dit dat de odds op vertrouwen in onbekenden 17% groter is wanneer iemand vrouw is. Dit is echter geen significant effect. Verder is het effect van leeftijd positief en significant ($b=0.03$; $p=0.000$). Voor elke stijging in leeftijd stijgen de odds met 0.03. Ook zijn positieve en significante effecten gevonden voor zowel respondenten met een middelhoog als hoog opleidingsniveau. Het effect van een middelhoog opgeleide respondent is redelijk groot ($b=0.367$; $p=0.016$). De odds-ratio is 1.44, wat betekent dat de odds op vertrouwen in onbekenden met 44% toeneemt wanneer iemand middelhoog opgeleid is ten opzichte van wanneer iemand laagopgeleid is. Ook het effect van een hoogopgeleide respondent is groot ($b=0.779$; $p=0.000$). De odds-ratio is 2.18. Wanneer een respondent hoogopgeleid is, stijgen de odds met 118%.

Vervolgens is de tweede analyse uitgevoerd. Dit is een logistische regressieanalyse met daarin het opleidingsniveau van zowel de vader als moeder samen met de controlevariabelen geslacht, leeftijd en het opleidingsniveau van de respondent. Vertrouwen in onbekenden is in dit model de afhankelijke variabele. Met deze analyse toetsen we of men meer vertrouwen heeft in onbekenden naarmate men ouders heeft die hoogopgeleid zijn.

Het effect van een middelhoog opgeleide vader is positief ($b=0.17$; $p=0.346$). De bijbehorende odds-ratio is 1.19, wat betekent dat de odds op vertrouwen in onbekenden met 19% toeneemt wanneer iemand een middelhoog opgeleide vader heeft ten opzichte van wanneer iemand een laagopgeleide vader heeft. Verder is ook het effect van een hoogopgeleide vader positief ($b=0.09$; $p=0.625$). Hierbij hoort een odds-ratio van 1.10.

Wanneer iemand een hoogopgeleide vader heeft, stijgt de odds op vertrouwen in onbekenden met 10% ten opzichte van wanneer iemand een laagopgeleide vader heeft. Beide effecten zijn redelijk klein en niet significant.

Daarnaast is het effect van een middelhoog opgeleide moeder positief ($b=0.18$; $p=0.374$). Er geldt een odds-ratio van 1.19, wat betekent dat de odds op vertrouwen in onbekenden stijgt met 19% wanneer iemand een middelhoog opgeleide moeder heeft ten opzichte van wanneer iemand een laagopgeleide moeder heeft. Verder is ook het effect van een hoogopgeleide moeder positief ($b=0.26$; $p=0.225$) met een odds-ratio van 1.30. Wanneer iemand een hoogopgeleide moeder heeft, stijgt de odds dat iemand vertrouwen heeft in onbekenden met 30% ten opzichte van wanneer iemand een laagopgeleide moeder heeft. Ook hier geldt dat de effecten niet significant zijn.

Omdat voor zowel hoogopgeleide vaders als hoogopgeleide moeders geen significant effect gevonden is, kunnen we de hypothese die stelt dat naarmate men ouders heeft die hoogopgeleid zijn men meer vertrouwen heeft in onbekenden niet aannemen.

Vervolgens is de derde analyse uitgevoerd. Dit is een lineaire regressieanalyse met daarin het opleidingsniveau van zowel de vader als moeder als onafhankelijke variabelen samen met de controlevariabelen geslacht, leeftijd en het opleidingsniveau van de respondent, maar dit keer met de variabele economische situatie gezin als afhankelijke variabele. Deze analyse is uitgevoerd om te kijken of gezinnen met hoogopgeleide ouders vroeger minder moeite hadden om rond te komen dan gezinnen met laagopgeleide ouders.

Het effect van een middelhoog opgeleide vader is positief en significant ($b=0.46$; $p=0.000$). Wanneer iemand een middelhoog opgeleide vader heeft, had het gezin vroeger minder moeite met rondkomen. Ook het effect van een hoogopgeleide vader is positief en significant ($b=0.55$; $p=0.000$), wat betekent dat iemand waarbij de vader hoogopgeleid is het gezin vroeger minder moeite had om rond te komen. De effecten zijn redelijk groot en

significant. Dit betekent dat gezinnen met een middelhoog- of hoogopgeleide over het algemeen minder moeite hadden met rondkomen dan gezinnen met een laagopgeleide vader.

Daarnaast is het effect van een middelhoog opgeleide moeder ook positief ($b=0.15$; $p=0.075$). Gezinnen waarbij de moeder middelhoog opgeleid is hadden vroeger minder moeite om rond te komen. Dit is een redelijk klein en niet significant effect. Het effect van een hoogopgeleide moeder op moeite met rondkomen is negatief en nauwelijks aanwezig ($b=-0.004$; $p=0.959$). De negatieve helling wijst erop dat hoogopgeleide moeders meer moeite hadden met rondkomen dan laagopgeleide moeders, maar dit is geen significant verschil.

Voor hoogopgeleide vaders is een redelijk groot en significant effect gevonden, terwijl voor hoogopgeleide moeders een klein en niet significant effect is gevonden. We kunnen daarom aannemen dat gezinnen met hoogopgeleide vaders vroeger over het algemeen geen moeite hadden met rondkomen, maar we kunnen niet aannemen dat gezinnen met hoogopgeleide moeders vroeger over het algemeen geen moeite hadden met rondkomen.

Ten slotte is de vierde en laatste analyse uitgevoerd. Het model dat hoort bij deze analyse is identiek aan het eerste model, maar de variabelen die horen bij de variabele economische situatie gezin zijn hier toegevoegd. Het gaat hierbij om de variabelen moeite met rondkomen, enigszins moeite met rondkomen en beetje moeite met rondkomen. Door deze analyse uit te voeren kan precies worden gekeken wat de economische situatie van het gezin bijdraagt aan de invloed van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden. De hypothese die getoetst wordt aan de hand van deze analyse is of het positieve effect van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden gedeeltelijk wordt verklaard door de economische situatie van het gezin.

Allereerst is gekeken naar vaders met een middelhoog- of hoog opleidingsniveau. De helling van een vader met een middelhoog opleidingsniveau ten opzichte van een vader met

een laag opleidingsniveau verandert van 0.17 in het eerste model naar 0.13 ($p=0.501$) in het derde model. De helling van een vader met een hoog opleidingsniveau ten opzichte van een vader met een laag opleidingsniveau verandert van 0.09 in het eerste model naar 0.02 ($p=0.898$) in het derde model. Bij zowel middelhoog- als hoogopgeleide vaders wordt het effect kleiner wanneer de variabele economische situatie gezin aan het model wordt toegevoegd. Toch is het effect niet erg groot en de hellingen zijn niet significant.

Vervolgens is gekeken naar moeders met een middelhoog- of hoog opleidingsniveau. De helling van een moeder met een middelhoog opleidingsniveau ten opzichte van een moeder met een laag opleidingsniveau verandert van 0.18 in het eerste model naar 0.17 ($p=0.391$) in het derde model. De helling van een moeder met een hoog opleidingsniveau ten opzichte van een moeder met een laag opleidingsniveau verandert van 0.26 in het eerste model naar 0.27 ($p=0.216$) in het derde model. Bij middelhoog opgeleide moeders wordt het effect iets kleiner, maar dit is geen significant verschil. Bij hoogopgeleide moeders stijgt de helling, maar ook dit verschil is niet significant.

Voor hoogopgeleide vaders en moeders zijn kleine en niet significante effecten gevonden voor het model waaraan de mediatie variabele economische situatie van het gezin is toegevoegd. De effecten lijken over het algemeen iets kleiner te worden met de extra variabele in het model, maar omdat er geen significante effecten gevonden zijn, kunnen we niet aannemen dat er sprake is van een mediërend effect. We nemen daarom niet aan dat het positieve effect van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden gedeeltelijk wordt verklaard door de economische situatie van het gezin.

Conclusie

In dit onderzoek is gekeken naar de invloed van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden. De onderzoeksvraag luidt als volgt: *Hebben mensen wiens ouders hoogopgeleid zijn meer vertrouwen in onbekenden dan mensen wiens ouders laagopgeleid zijn?* Het is belangrijk om onderzoek te doen naar vertrouwen in onbekenden, omdat het een rol speelt in zowel het dagelijks leven van individuen als op maatschappelijk niveau. Zo kan vertrouwen op individueel niveau zorgen voor meer levenstevredenheid (Welch et al., 2005) en kan het op maatschappelijk niveau leiden tot meer groei en efficiëntie (Delhey & Newton, 2003). Omdat ouders een grote rol spelen in de vroege jeugd van hun kinderen, is het goed om te kijken naar hun rol in het overbrengen van vertrouwen aan hun kinderen. Er kan daarbij worden gekeken of het opleidingsniveau van ouders hierbij een verschil maakt, zodat dit kan worden gebruikt om het vertrouwen in onbekenden onder de gehele bevolking te maximaliseren.

Om de onderzoeksvraag te beantwoorden zijn drie hypothesen opgesteld. De eerste hypothese stelt dat men meer vertrouwen heeft in onbekenden wanneer men ouders heeft die hoogopgeleid zijn. De tweede hypothese stelt dat gezinnen waarbij de ouders hoogopgeleid zijn minder moeite hadden om rond te komen vroeger dan gezinnen waarbij de ouders laagopgeleid zijn. De derde hypothese stelt dat het positieve effect van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden gedeeltelijk wordt verklaard door de economische situatie van het gezin.

Voor de eerste hypothese is geen ondersteuning gevonden. De verbanden tussen het opleidingsniveau van zowel de vader als de moeder zijn positief, maar er is geen bewijs gevonden dat een hoog opleidingsniveau van iemands ouders leidt tot meer vertrouwen in onbekenden. Deze uitspraak kan daarom niet worden gegeneraliseerd naar de populatie. Dit is niet in lijn met de socialisatieverklaring. Hoogopgeleiden hebben een optimistische kijk

(O'Connor & Graham, 2019), waardoor het aannemelijk is dat ze meer vertrouwen hebben in onbekenden. De verwachting was dat hoogopgeleide ouders deze optimistische kijk zouden doorgeven aan hun kind, omdat zij deel uitmaken van het informele netwerk van het kind en een grote rol spelen in de vroege jeugd. Deze optimistische zou er dan voor zorgen dat de kinderen zelf ook meer vertrouwen hebben, onder andere in onbekenden.

Voor de tweede hypothese is deels ondersteuning gevonden. Er is namelijk bewijs gevonden voor de aanname dat gezinnen met een hoogopgeleide vader vroeger minder moeite hadden om rond te komen. Dit wijst erop dat de economische situatie van het gezin beter is wanneer de vader hoogopgeleid is. Dit is in overeenstemming met wat we vooraf verwachten, omdat uit de literatuur blijkt dat hoogopgeleide Nederlanders meer verdienen dan laagopgeleide Nederlanders (Centraal Bureau voor de Statistiek, 2011; Zwart, 2021) en dus minder moeite hebben om rond te komen.

Daar tegenover staat dat er geen bewijs is gevonden voor de aanname dat gezinnen met een hoogopgeleide moeder vroeger minder moeite hadden om rond te komen. Er lijkt zelfs sprake te zijn van een negatief verband, wat betekent dat gezinnen met een hoogopgeleide moeder juist meer moeite hadden met rondkomen dan gezinnen met een laagopgeleide moeder.

Een oorzaak voor dit negatieve verband kan te maken hebben met de tijd waarin de moeders leefden. De gemiddelde leeftijd van respondenten in de steekproef is 53 jaar, wat betekent dat veel moeders van respondenten geboren zijn vóór of rond de Tweede Wereldoorlog. Vrouwen stopten vaak met werken wanneer ze kinderen kregen, terwijl hun mannelijke partners carrière maakten (Strader, 2021). Daarnaast verdienden de vrouwen die wel bleven werken minder dan mannen, een verschijnsel dat de *gender wage gap* wordt genoemd (Strader, 2021). Dat hoogopgeleide moeders vaker stopten met werken en dat de vrouwen die werkten minder verdienden dan mannen, kan deels verklaren waarom

hoogopgeleide moeders – in tegenstelling tot hoogopgeleide vaders – vroeger moeite hadden met rondkomen. Voor dit negatieve verband tussen het opleidingsniveau van moeders en moeite hebben om rond te komen is echter geen bewijs gevonden en het kan daarom niet worden gegeneraliseerd naar de populatie.

Voor de derde hypothese is wederom geen ondersteuning gevonden. Het positieve effect van het opleidingsniveau van iemands ouders op de mate van vertrouwen in onbekenden lijkt gedeeltelijk te worden verklaard door de economische situatie van het gezin. Toch is er geen bewijs gevonden voor deze aanname, waardoor het niet kan worden gegeneraliseerd naar de populatie.

Dit is niet in overeenstemming met wat we vooraf verwachtten. Enerzijds verwachtten we aan de hand van de welvaartsverklaring dat een hoog opleidingsniveau van ouders er via de economische situatie van het gezin voor zou zorgen dat kinderen een groter financieel vangnet zouden hebben. Dit zou ertoe leiden dat een deel van de risico's van het vertrouwen van onbekenden zou worden opgevangen. Anderzijds verwachtten we aan de hand van de buurteffect verklaring dat een hoog opleidingsniveau van ouders er via de economische situatie van het gezin voor zou zorgen dat kinderen opgroeien in een veiligere buurt, waardoor kinderen meer vertrouwen zouden hebben in buurtgenoten. Door middel van het spill-over effect zouden deze jongeren dan meer vertrouwen hebben in onbekenden.

Concluderend kunnen we stellen dat mensen wiens ouders hoogopgeleid zijn niet noodzakelijk meer vertrouwen hebben in onbekenden dan mensen wiens ouders laagopgeleid zijn. In de volgende paragraaf zullen een aantal beperkingen van het onderzoek worden toegelicht en zal enkele aanbevelingen worden gedaan voor toekomstig onderzoek.

Discussie

Dit onderzoek kent een aantal beperkingen waarmee rekening moet worden gehouden. De grootste beperking is dat het onderzoek zich grotendeels richt op de ouders van respondenten, maar dat deze ouders zelf niet worden ondervraagd. De respondenten beantwoorden vragen over het opleidingsniveau van hun ouders en over de economische situatie van het gezin vroeger. Omdat veel respondenten zelf al wat ouder zijn ($y=53$) is de kans aanwezig dat respondenten zich het opleidingsniveau niet goed meer kunnen herinneren. Het is ten slotte een tijd geleden dat hun ouders een opleidingsniveau hebben afgerond en het is aannemelijk dat deze informatie is weggezaakt bij hun kinderen. Bovendien veranderen schoolsystemen nog weleens, waardoor het kan voorkomen dat een respondent zich vergist in het opleidingsniveau van zijn of haar vader en/of moeder.

Daarnaast is de kans aanwezig dat de respondenten zich de economische situatie van het gezin vroeger niet goed herinneren of een onjuist beeld hebben van de situatie. Respondenten waren zelf kinderen toen ze bij hun ouders woonden en het kan daarom lastig zijn om een inschatting te maken van de economische situatie van het gezin vroeger. Niet alle ouders zijn transparant en ouders kunnen het voor hun kinderen hebben verzwegen wanneer ze het financieel moeilijk hadden. Gezinnen die in grote huizen woonden en waarbij het de kinderen aan niets ontbrak kunnen het - net als bij gezinnen waarbij het voor de kinderen duidelijk was dat de ouders niet goed rond konden komen - financieel moeilijk hebben gehad. Dat respondenten de vraag over de economische situatie van het gezin vroeger zelf hebben beantwoord in plaats van de ouders, maakt de meting niet erg objectief en betrouwbaar.

Een andere beperking binnen het onderzoek is dat de onderzoeksvraag niet precies genoeg kon worden beantwoord. De onderzoeksvraag gaat namelijk over de invloed van het opleidingsniveau van iemands ouders, maar het opleidingsniveau van de vader en moeder zijn onafhankelijk van elkaar bekeken. Hiermee kunnen we wel iets zeggen over de invloed van

een hoog opleidingsniveau van de vader en de invloed van een hoog opleidingsniveau van de moeder, maar niet over de invloed van het hebben van zowel een vader als moeder met een hoog opleidingsniveau.

Verder is de vraag in hoeverre gezinnen vroeger moeite hadden met rondkomen niet naar wens bevraagd. Respondenten konden namelijk kiezen uit vier antwoorden waarbij de categorieën dicht bij elkaar liggen. Zo konden respondenten bijvoorbeeld kiezen voor 'enigszins moeite met rondkomen' of een 'beetje moeite met rondkomen'. De grens tussen 'enigszins' en 'beetje' is erg subjectief en respondenten kunnen de vragen hierdoor anders hebben beantwoord, terwijl ze misschien hetzelfde bedoelden.

Ditzelfde geldt voor de vraag over het vertrouwen in onbekenden. Ook hier kon worden gekozen uit vier antwoorden. Hoewel de vragen duidelijk zijn gespecificeerd, zien we dat de meeste respondenten kiezen voor 'vertrouw ik een beetje' of 'vertrouw ik niet heel erg' en dat minder mensen kiezen voor 'vertrouw ik helemaal' of 'vertrouw ik helemaal niet'.

Naast een aantal inhoudelijke beperkingen van het onderzoek, is er een implicatie die moet worden benoemd. Een van de belangrijkste verwachtingen van het onderzoek was dat mensen met hoogopgeleide ouders meer vertrouwen hebben in onbekenden dan mensen met laagopgeleide ouders. Hiervoor werd geen ondersteuning gevonden. Uit de theorie is de socialisatieverklaring afgeleid dat meer optimisme leidt tot meer vertrouwen in onbekenden en dat hoogopgeleiden een sterke band hebben met hun kinderen, waardoor het vertrouwen kan worden overgedragen van de ouders op de kinderen. Er is hierbij te weinig ingegaan op hoe vertrouwen dan precies van persoon tot persoon kan worden overgedragen en daarom is het goed om hier meer (literatuur)onderzoek naar te doen. Daarbij kan in vragenlijsten een onderdeel worden toegevoegd over bijvoorbeeld opvoeding, inclusief vragen of en hoe ouders vertrouwen overdragen aan hun kinderen.

Uit bovenstaande beperkingen en implicatie van het onderzoek volgt een tweetal adviezen. Het eerste advies is om een getrapte steekproef uit te voeren. Dit houdt in dat naast de respondenten, ook de ouders worden bevraagd. Op deze manier kan meer informatie worden verkregen over de economische situatie van het gezin en over hoeveel vertrouwen ouders hebben in onbekenden. Daarbij kan ook worden gevraagd naar onderwerpen over opvoeding en woonsituatie, zodat kan worden achterhaald of vertrouwen iets is wat kan worden doorgegeven aan kinderen.

Het tweede advies is om de vragen over de economische situatie van het gezin en de mate van vertrouwen in onbekenden anders te stellen, in verband met de inhoudsvaliditeit. Voor de vraag over de economische situatie kan bijvoorbeeld een vraag worden toegevoegd die ingaat op het daadwerkelijke inkomen van de ouders. Op die manier kan de economische situatie van het gezin worden gecategoriseerd op basis van het gezamenlijk inkomen van gezinnen. Zo komt een objectievere meting tot stand. Voor de vraag over het vertrouwen in onbekenden is het goed om te kijken of een tien-puntsschaal kan worden gemaakt, zodat nog beter kan worden gekeken hoeveel vertrouwen respondenten hebben.

Literatuur

- Centraal Bureau voor de Statistiek. (2011, 23 maart). *Inkomen hoogopgeleiden bijna het dubbele van dat van laagopgeleiden* <https://www.cbs.nl/nlnl/nieuws/2011/12/inkomen-hoogopgeleiden-bijna-het-dubbele-van-dat-van-laagopgeleiden>.
- Centraal Bureau voor de Statistiek. (2023, 30 mei). *Opleidingsniveau*. <https://www.cbs.nl/nlnl/nieuws/2019/33/verschil-levensverwachting-hoog-en-laagopgeleid-groeit/opleidingsniveau>
- Delhey, J., & Newton, K. (2003). Who trusts?: The origins of social trust in seven societies. *European Societies*, 5(2), 93-137.
- De Vries, A. P. R. (2021). 1. Inleiding. *Centraal Bureau voor de Statistiek*. <https://www.cbs.nl/nl-nl/longread/discussion-papers/2021/invulling-praktisch-en-theoretisch-opgeleiden/1-inleiding>
- EVS (2022). EVS Trend File 1981-2017 – Sensitive Dataset. GESIS, Cologne. ZA7504 Data file Version 2.0.0, <https://doi.org/10.4232/1.14022>
- Farrington, D. P. (2017). How and why do Amherst and Newark differ in crime rates? Comments on “America’s Safest City”. *Crime Law and Social Change*. <https://doi.org/10.1007/s10611-017-9683-5>
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, 116(3), 429–456. <https://doi.org/10.1037/0033-2909.116.3.429>
- Galster, G. (2002). An economic efficiency analysis of deconcentrating poverty populations. *Journal of Housing Economics*, 11, 303–329.
- Hooghe, M. (2007). Social Capital and Diversity Generalized Trust, Social Cohesion and Regimes of Diversity. *Canadian Journal of Political Science/Revue Canadienne De Science Politique*, 40(3), 709-732. doi:10.1017/S0008423907070722
- Hooghe, M., Marien, S., & De Vroome, T. (2012). The cognitive basis of trust. The relation between education, cognitive ability, and generalized and political trust. *Intelligence*, 40(6), 604–613. <https://doi.org/10.1016/j.intell.2012.08.006>
- Leenheer, S., Gesthuizen, M., & Savelkoul, M. (2021). Two-Way, One-Way or Dead-End Streets? Financial and Social Causes and Consequences of Generalized Trust. *Social Indicators Research*, 155(3), 915-937. <https://doi.org/10.1007/s11205-020-02591-6>.

- Rontos, K., & Roumeliotou, M. (2013). Generalized social trust in Greece and its association with demographic and socio-economic predictors. *Portuguese Journal of Social Science*, 12(1), 63–84. https://doi.org/10.1386/pjss.12.1.63_1
- O'Connor, K., & Graham, C. (2019). Longer, more optimistic, lives: Historic optimism and life expectancy in the United States. *Journal of Economic Behavior and Organization*, 168, 374–392. <https://doi.org/10.1016/j.jebo.2019.10.018>
- Uslaner, E. M. (2002). *The moral foundations of trust*. Cambridge, UK: Cambridge University Press.
- Van Ingen, E., & Bekkers, R. (2015). Generalized trust through civic engagement? Evidence from five national panel studies. *Political Psychology*, 35(3), 277–294.
- Uslaner, E. M. (1998). Social Capital, Television, and the “Mean World”: Trust, Optimism, and Civic Participation. *Political Psychology*, 19(3), 441–467. <https://doi.org/10.1111/0162-895x.00113>
- Strader, E. (2022). State work-family contexts and the wage gap by gender and parenthood. *Family Relations*, 71(4), 1713–1730. <https://doi-org.proxy-ub.rug.nl/10.1111/fare.12680>
- Van Ham, M., Boschman, S., & Vogel, M. (2018). Incorporating Neighborhood Choice in a Model of Neighborhood Effects on Income. *Demography*, 55(3), 1069–1090. <https://doi.org/10.1007/s13524-018-0672-9>
- Woolcock, M. (1998). Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework. *Theory and Society*, 27(2), 151–208. <http://www.jstor.org/stable/657866>
- Yamagishi, T., Kikuchi, M., & Kosugi, M. (1999). Trust, Gullibility, and Social Intelligence. *Asian Journal of Social Psychology*, 2(1), 145–161. <https://doi.org/10.1111/1467-839x.00030>
- Zwart, C. (2023, 26 april). *Modaal en gemiddeld inkomen per leeftijd, opleidingsniveau en branche*. <https://bieb.knab.nl/inkomsten-uitgaven/modaal-en-gemiddeld-inkomen-per-leeftijd-opleidingsniveau-en-branche#opleidingsniveau>

Bijlage 1

Selectie uit de steekproef

In het onderzoek wordt gekeken naar data uit jaar 2017. Om die reden is allereerst de bewerking uitvoert waarbij alle andere jaartallen uit het bestand zijn geselecteerd. De syntax voor deze bewerking is als volgt:

```
SELECT IF (year = 2017).  
EXECUTE.
```

Omdat het onderzoek specifiek gaat over Nederland is er daarna opnieuw een selectie gemaakt. Dit keer zijn participanten uit andere landen dan uit Nederland uit de steekproef geselecteerd. De syntax voor deze bewerking is als volgt:

```
SELECT IF (country = 528).  
EXECUTE.
```

De oorspronkelijke variabelen worden inclusief missende data bekeken. Na de bewerkingen van de variabelen wordt er slechts gekeken naar respondenten die een score hebben ingevuld voor alle variabelen. De rest wordt uit de data geselecteerd en daardoor niet meegenomen in de beschrijvende statistieken.

Allereerst is er een regressiemodel gerund met daarin alle variabele die moeten worden meegenomen. Daarvan zijn de ongestandaardiseerde residuen opgeslagen. Vervolgens is er een dummyvariabele aangemaakt (obs) en is er een filter aangemaakt. Door dit te doen kunnen de beschrijvende statistieken worden uitgerekend voor alle complete cases. De syntax is hieronder weergegeven:

```
LOGISTIC REGRESSION VARIABLES Vertrouwen_r  
/METHOD=ENTER age geslacht_r v270 v243_r v262_r v263_r  
/SAVE=RESID  
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

```
RECODE RES_1 (SYSMIS=0) (ELSE=1) INTO obs.  
EXECUTE.
```



```

USE ALL.
COMPUTE filter_$=(obs = 1).
VARIABLE LABELS filter_$ 'obs = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.

```

Beschrijvende statistieken

Vertrouwen in onbekenden

Om de beschrijvende statistieken zoals het gemiddelde, de standaarddeviatie en de verdeling van de oorspronkelijke variabele vertrouwen in onbekenden te bekijken is de volgende syntax gerund:

```

FREQUENCIES VARIABLES=v35
  /STATISTICS=STDDEV MINIMUM
  MAXIMUM MEAN MEDIAN MODE
  /ORDER=ANALYSIS.

```

Hieruit is de volgende output gekomen:

Statistics		
how much you trust: people you		
N	Valid	2338
	Missing	66
Mean		2,19
Median		2,00
Mode		2
Std. Deviation		,629
Minimum		1
Maximum		4

how much you trust: people you meet for the first time (Q8D)					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	trust completely	183	7,6	7,8	7,8
	trust somewhat	1615	67,2	69,1	76,9
	do not trust very much	443	18,4	18,9	95,9
	do not trust at all	97	4,0	4,1	100,0
	Total	2338	97,3	100,0	
Missing	no answer	7	,3		
	dont know	59	2,5		
	Total	66	2,7		
Total		2404	100,0		

B1.1: beschrijvende statistieken van vertrouwen (onbewerkt)

Omdat de variabele bestaat uit categorieën die goed samengevoegd kunnen worden, is de variabele gehercodeerd naar een dichotome variabele. De syntax staat hier onder weergegeven:

```

RECODE v35 (1=1) (2=1) (3=0) (4=0) (-10 thru -1=SYSMIS) INTO
  Vertrouwen_r.
EXECUTE.

```

Vervolgens is voor de gehercodeerde variabele van vertrouwen in onbekenden opnieuw gekeken naar de beschrijvende statistieken. Dit keer zijn alleen de respondenten meegenomen die een score hebben ingevuld op elk van de variabelen (door middel van het filter met obs). De syntax en de output staan hieronder weergegeven.

```
FREQUENCIES VARIABLES=Vertrouwen_r
/STATISTICS=STDDEV MINIMUM
MAXIMUM MEAN MEDIAN MODE
/ORDER=ANALYSIS.
```

Statistics

Vertrouwen_r

N	Valid	1968
	Missing	0
Mean		,7840
Median		1,0000
Mode		1,00
Std. Deviation		,41159
Minimum		,00
Maximum		1,00

Vertrouwen_r

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	425	21,6	21,6	21,6
	1,00	1543	78,4	78,4	100,0
Total		1968	100,0	100,0	

B1.2: beschrijvende statistieken van vertrouwen (bewerkt)

Opleidingsniveau vader

Om de beschrijvende statistieken zoals het gemiddelde, de standaarddeviatie en de verdeling van de oorspronkelijke variabele opleidingsniveau vader te bekijken is de volgende syntax gerund:

```
FREQUENCIES VARIABLES=v262_r
/STATISTICS=STDDEV MINIMUM
MAXIMUM MEAN MEDIAN MODE
/ORDER=ANALYSIS.
```

Hieruit is de volgende output gekomen:

Statistics

educational level father: recodec

N	Valid	2098
	Missing	306
Mean		1,57
Median		1,00
Mode		1
Std. Deviation		,816
Minimum		1
Maximum		3

educational level father: recoded (Q99)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	lower	1350	56,2	64,3	64,3
	medium	308	12,8	14,7	79,0
	higher	440	18,3	21,0	100,0
	Total	2098	87,3	100,0	
Missing	no answer	37	1,5		
	dont know	269	11,2		
	Total	306	12,7		
Total		2404	100,0		

B1.3: beschrijvende statistieken van opleidingsniveau vader (onbewerkt)

Omdat opleidingsniveau vader drie categorieën heeft, zijn er twee dummyvariabelen gemaakt. De syntax is als volgt:

```
RECODE v262_r (1=0) (2=1) (3=0) (-10 thru -1=SYSMIS) INTO OpleidingVader_M.
RECODE v262_r (1=0) (2=0) (3=1) (-10 thru -1=SYSMIS) INTO OpleidingVader_H.
EXECUTE.
```

Vervolgens zijn nogmaals de beschrijvende statistieken uitgevoerd, dit keer door alleen respondenten mee te nemen die een score hebben ingevuld voor alle variabelen. Er zijn dus geen missende waarden. De syntax en output staan hieronder weergegeven:

```
FREQUENCIES VARIABLES=v262_r
/STATISTICS=STDDEV MINIMUM
MAXIMUM MEAN MEDIAN MODE
/ORDER=ANALYSIS.
```

Statistics

educational level father: recoded (Q99)

N	Valid	1968
	Missing	0
Mean		1,57
Median		1,00
Mode		1
Std. Deviation		,820
Minimum		1
Maximum		3

educational level father: recoded (Q99)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	lower	1265	64,3	64,3	64,3
	medium	283	14,4	14,4	78,7
	higher	420	21,3	21,3	100,0
	Total	1968	100,0	100,0	

B1.4: beschrijvende statistieken van opleidingsniveau vader (bewerkt)

Opleidingsniveau moeder

Om de beschrijvende statistieken zoals het gemiddelde, de standaarddeviatie en de verdeling van de oorspronkelijke variabele opleidingsniveau moeder te bekijken is de volgende syntax gerund:

```
FREQUENCIES VARIABLES=v263_r
/STATISTICS=STDDEV MINIMUM
MAXIMUM MEAN MEDIAN MODE
/ORDER=ANALYSIS.
```

Hieruit is de volgende output gekomen:

Statistics

educational level mother: recode

N	Valid	2127
	Missing	277
Mean		1,39
Median		1,00
Mode		1
Std. Deviation		,708
Minimum		1
Maximum		3

educational level mother: recoded (Q100)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	lower	1574	65,5	74,0	74,0
	medium	274	11,4	12,9	86,9
	higher	279	11,6	13,1	100,0
	Total	2127	88,5	100,0	
Missing	no answer	38	1,6		
	dont know	239	9,9		
	Total	277	11,5		
Total		2404	100,0		

B1.5: beschrijvende statistieken van opleidingsniveau moeder (onbewerkt)

Omdat opleidingsniveau moeder drie categorieën heeft, zijn er twee dummyvariabelen gemaakt. De syntax is als volgt:

```
RECODE v262_r (1=0) (2=1) (3=0) (-10 thru -1=SYSMIS) INTO OpleidingMoeder_M.
RECODE v262_r (1=0) (2=0) (3=1) (-10 thru -1=SYSMIS) INTO OpleidingMoeder_H.
EXECUTE.
```

Vervolgens zijn nogmaals de beschrijvende statistieken uitgevoerd, dit keer door alleen respondenten mee te nemen die een score hebben ingevuld voor alle variabelen. Er zijn dus geen missende waarden. De syntax en output staan hieronder weergegeven:

```
FREQUENCIES VARIABLES=v262_r
/STATISTICS=STDDEV MINIMUM
MAXIMUM MEAN MEDIAN MODE
/ORDER=ANALYSIS.
```

Statistics

educational level mother: recoded (Q100)

N	Valid	1968
	Missing	0
Mean		1,39
Median		1,00
Mode		1
Std. Deviation		,711
Minimum		1
Maximum		3

educational level mother: recoded (Q100)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	lower	1456	74,0	74,0	74,0
	medium	250	12,7	12,7	86,7
	higher	262	13,3	13,3	100,0
	Total	1968	100,0	100,0	

B1.6: beschrijvende statistieken van opleidingsniveau moeder (bewerkt)

Economische situatie gezin

Om de beschrijvende statistieken zoals het gemiddelde, de standaarddeviatie en de verdeling van de oorspronkelijke variabele economische situatie gezin te bekijken is de volgende syntax gerund:

```
FREQUENCIES VARIABLES=v270
  /STATISTICS=STDDEV MINIMUM
  MAXIMUM MEAN MEDIAN MODE
  /ORDER=ANALYSIS.
```

Hieruit is de volgende output gekomen:

Statistics		
at age 14, parent(s) had problem		
N	Valid	2322
	Missing	82
Mean		3,02
Median		3,00
Mode		4
Std. Deviation		1,115
Minimum		1
Maximum		6

at age 14, parent(s) had problems making ends meet (Q104D)					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	yes	317	13,2	13,7	13,7
	to some extent	457	19,0	19,7	33,3
	a little bit	416	17,3	17,9	51,2
	no	1128	46,9	48,6	99,8
	does not apply to me (spontaneous)	4	,2	,2	100,0
Total		2322	96,6	100,0	
Missing	no answer	17	,7		
	dont know	65	2,7		
	Total	82	3,4		
Total		2404	100,0		

BI.7: beschrijvende statistieken de economische situatie van het gezin (onbewerkt)

Van de variabele zijn 3 dummyvariabelen gemaakt. De syntax is als volgt:

```
RECODE v270 (1=1) (2=0) (3=0) (4=0) (-10 thru -1=SYSMIS) INTO
Moeite_Rondkomen_Ja.
RECODE v270 (1=0) (2=1) (3=0) (4=0) (-10 thru -1=SYSMIS) INTO
Moeite_Rondkomen_Enigsz.
RECODE v270 (1=0) (2=0) (3=1) (4=0) (-10 thru -1=SYSMIS) INTO
Moeite_Rondkomen_Beetje.
EXECUTE.
```

Vervolgens zijn nogmaals de beschrijvende statistieken uitgevoerd, dit keer door alleen respondenten mee te nemen die een score hebben ingevuld voor alle variabelen. Er zijn dus geen missende waarden. De syntax en output staan hieronder weergegeven:

```
FREQUENCIES VARIABLES=v270
  /STATISTICS=STDDEV MINIMUM
  MAXIMUM MEAN MEDIAN MODE
  /ORDER=ANALYSIS.
```

Statistics

at age 14, parent(s) had problems making ends meet (Q104D)

N	Valid	1968
		Missing
Mean		3,05
Median		3,00
Mode		4
Std. Deviation		1,094
Minimum		1
Maximum		4

at age 14, parent(s) had problems making ends meet (Q104D)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	yes	248	12,6	12,6	12,6
	to some extent	390	19,8	19,8	32,4
	a little bit	354	18,0	18,0	50,4
	no	976	49,6	49,6	100,0
	Total	1968	100,0	100,0	

B1.8: beschrijvende statistieken van de economische situatie van het gezin (bewerkt)

Geslacht

Om de beschrijvende statistieken zoals het gemiddelde, de standaarddeviatie en de verdeling van de oorspronkelijke variabele geslacht te bekijken is de volgende syntax gerund:

```
FREQUENCIES VARIABLES=v225
  /STATISTICS=STDDEV MINIMUM
  MAXIMUM MEAN MEDIAN MODE
  /ORDER=ANALYSIS.
```

Hieruit is de volgende output gekomen:

Statistics

sex respondent (Q63)

N	Valid	2404
		Missing
Mean		1,53
Median		2,00
Mode		2
Std. Deviation		,499
Minimum		1
Maximum		2

sex respondent (Q63)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	male	1120	46,6	46,6	46,6
	female	1284	53,4	53,4	100,0
	Total	2404	100,0	100,0	

B1.9: beschrijvende statistieken van geslacht (onbewerkt)

Er is voor gekozen om te hercoderen naar 0 = man en 1 = vrouw. De syntax voor deze hercodering is als volgt:

```
RECODE v225 (1=0) (2=1) (-10 thru -
1=SYSMIS) INTO Geslacht_r.
EXECUTE.
```

Vervolgens zijn opnieuw de beschrijvende statistieken uitgerekend, dit keer voor de gehercodeerde variabele. De syntax en output zijn hieronder weergegeven.

```

FREQUENCIES VARIABLES=Geslacht_r
/STATISTICS=STDDEV MINIMUM
MAXIMUM MEAN MEDIAN MODE
/ORDER=ANALYSIS.

```

Statistics

Geslacht_r		
N	Valid	1968
	Missing	0
Mean		,5279
Median		1,0000
Mode		1,00
Std. Deviation		,49935
Minimum		,00
Maximum		1,00

Geslacht_r					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	929	47,2	47,2	47,2
	1,00	1039	52,8	52,8	100,0
Total		1968	100,0	100,0	

B1.10: beschrijvende statistieken van geslacht (bewerkt)

Leeftijd

Om de beschrijvende statistieken zoals het gemiddelde, de standaarddeviatie en de verdeling van de oorspronkelijke variabele leeftijd te bekijken is de volgende syntax gerund:

```

FREQUENCIES VARIABLES=age
/STATISTICS=STDDEV MINIMUM
MAXIMUM MEAN MEDIAN MODE
/HISTOGRAM NORMAL
/ORDER=ANALYSIS.

```

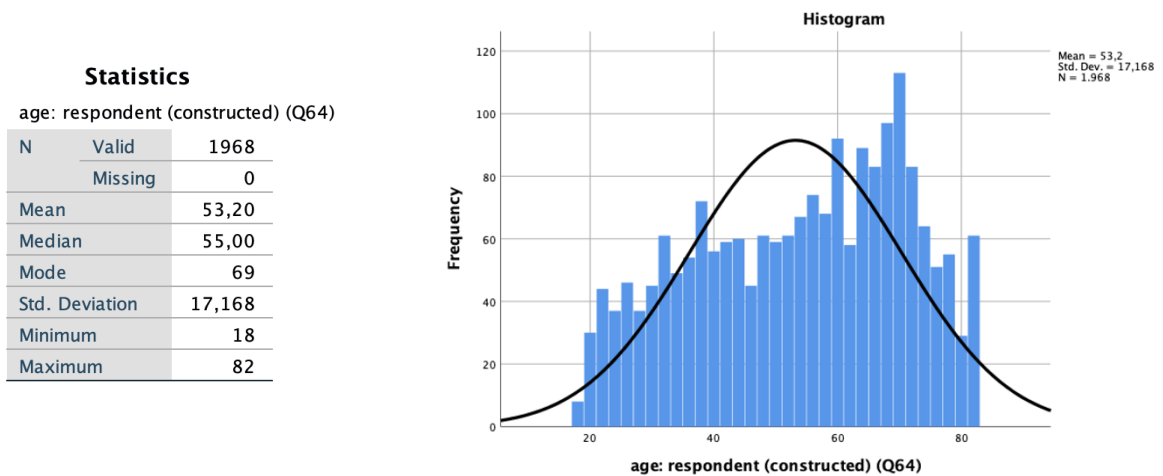
Hieruit is de volgende output gekomen:

Statistics		
age: respondent (constructed) (
N	Valid	2404
	Missing	0
Mean		53,02
Median		55,00
Mode		69
Std. Deviation		17,318
Minimum		18
Maximum		82



B1.11: beschrijvende statistieken van leeftijd (onbewerkt)

Er is geen bewerking uitgevoerd. Wel zijn opnieuw de beschrijvende statistieken uitgerekend voor de respondenten die een score hebben ingevuld op alle variabelen. De output is als volgt:



B1.12: beschrijvende statistieken van leeftijd (onbewerkt)

Opleidingsniveau respondent

Om de beschrijvende statistieken zoals het gemiddelde, de standaarddeviatie en de verdeling van de oorspronkelijke variabele opleidingsniveau respondent te bekijken is de volgende syntax gerund:

```
FREQUENCIES VARIABLES=v243_r
  /STATISTICS=STDDEV MINIMUM
  MAXIMUM MEAN MEDIAN MODE
  /ORDER=ANALYSIS.
```

Hieruit is de volgende output gekomen:

Statistics			educational level respondent: recoded (Q81)				
educational level respondent: re				Frequency	Percent	Valid Percent	Cumulative Percent
N	Valid	2369	Valid	lower	715	29,7	30,2
	Missing	35		medium	656	27,3	57,9
Mean		2,12		higher	998	41,5	100,0
Median		2,00	Total	Total	2369	98,5	100,0
Mode		3	Missing	no answer	15	,6	
Std. Deviation		,842		dont know	20	,8	
Minimum		1	Total	Total	35	1,5	
Maximum		3	Total	Total	2404	100,0	

B1.13: beschrijvende statistieken van opleidingsniveau respondent (onbewerkt)

Omdat opleidingsniveau respondent drie categorieën heeft, zijn er twee dummyvariabelen gemaakt. De syntax is als volgt:


```
RECODE v243_r (1=1) (2=0) (3=0) (-10 thru -1=SYSMIS) INTO OpleidingRes_M.
RECODE v243_r (1=0) (2=1) (3=0) (-10 thru -1=SYSMIS) INTO OpleidingRes_H.
EXECUTE.
```

Vervolgens zijn nogmaals de beschrijvende statistieken uitgevoerd, dit keer door alleen respondenten mee te nemen die een score hebben ingevuld voor alle variabelen. Er zijn dus geen missende waarden. De syntax en output staan hieronder weergegeven:

```
FREQUENCIES VARIABLES=v243_r
/STATISTICS=STDDEV MINIMUM
MAXIMUM MEAN MEDIAN MODE
/ORDER=ANALYSIS.
```

Statistics

educational level respondent: recoded (Q81)

N	Valid	1968
	Missing	0
Mean		2,17
Median		2,00
Mode		3
Std. Deviation		,836
Minimum		1
Maximum		3

educational level respondent: recoded (Q81)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	lower	546	27,7	27,7	27,7
	medium	533	27,1	27,1	54,8
	higher	889	45,2	45,2	100,0
	Total	1968	100,0	100,0	

B1.14: beschrijvende statistieken van opleidingsniveau respondent (bewerkt)

Bijlage 2

Associatiematen

In het onderzoek wordt gebruik gemaakt van één continue variabele, twee binaire categorische variabelen en vier variabelen met meer dan twee categorieën.

Allereerst is gekeken naar de correlatie tussen de continue variabele (leeftijd) en de binaire categorische variabelen (vertrouwen in onbekenden en geslacht). Om tot een correlatie te komen is de volgende syntax gerund. Deze is hieronder weergegeven:

```
CORRELATIONS
/VARIABLES=age
Vertrouwen_r Geslacht_r
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

De bijbehorende output is als volgt:

		age: respondent (constructed) (Q64)	Vertrouwen_r	Geslacht_r
age: respondent (constructed) (Q64)	Pearson Correlation	1	,144**	-,047*
	Sig. (2-tailed)		,000	,037
	N	1968	1968	1968
Vertrouwen_r	Pearson Correlation	,144**	1	,021
	Sig. (2-tailed)	,000		,358
	N	1968	1968	1968
Geslacht_r	Pearson Correlation	-,047*	,021	1
	Sig. (2-tailed)	,037	,358	
	N	1968	1968	1968

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

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

B2.1: correlatie tussen leeftijd en vertrouwen in onbekenden en geslacht

De samenhang tussen de continue variabele (leeftijd) en de categorische variabelen (opleidingsniveau vader, opleidingsniveau moeder, economische situatie gezin en opleidingsniveau respondent) is bekeken door regressieanalyses uit te voeren. Hierbij is gekeken naar de kolom R onder Model Summary.

De syntax en bijbehorende output is als volgt:

```

REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
CHANGE
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT age
/METHOD=ENTER v262_r.

```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			
						F Change	df1	df2	Sig. F Change
1	,301 ^a	,090	,090	16,379	,090	195,265	1	1966	,000

a. Predictors: (Constant), educational level father: recoded (Q99)

B2.2: correlatie tussen leeftijd en opleidingsniveau vader

```

REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
CHANGE
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT age
/METHOD=ENTER v263_r.

```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			
						F Change	df1	df2	Sig. F Change
1	,404 ^a	,164	,163	15,706	,164	384,276	1	1966	,000

a. Predictors: (Constant), educational level mother: recoded (Q100)

B2.3: correlatie tussen leeftijd en opleidingsniveau moeder

```

REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
CHANGE
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT age
/METHOD=ENTER v270.

```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			
						F Change	df1	df2	Sig. F Change
1	,199 ^a	,040	,039	16,830	,040	80,907	1	1966	,000

a. Predictors: (Constant), at age 14, parent(s) had problems making ends meet (Q104D)

B2.3: correlatie tussen leeftijd en economische situatie gezin

```

REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
CHANGE
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT age
/METHOD=ENTER v243_r.

```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			
						F Change	df1	df2	Sig. F Change
1	,249 ^a	,062	,061	16,633	,062	129,634	1	1966	,000

a. Predictors: (Constant), educational level respondent: recoded (Q81)

B2.4: correlatie tussen leeftijd en opleidingsniveau respondent

De samenhang tussen categorische variabelen is bekeken door de Cramer's V uit te rekenen. Omdat deze methode niets zegt over de richting van de samenhang is deze daarom bekeken aan de hand van kruistabellen.

Eerst zijn kruistabellen bekeken tussen vertrouwen in onbekenden en de overige categorische variabelen. De syntax is als volgt:

CROSSTABS

```
/TABLES=Vertrouwen_r BY v262_r v263_r  
v270 Geslacht_r v243_r  
/FORMAT=AVALUE TABLES  
/STATISTICS=PHI  
/CELLS=COUNT  
/COUNT ROUND CELL.
```

Hierbij hoort de volgende output:

Vertrouwen_r * educational level father: recoded (Q99)

Count		Crosstab			Total
		educational level father: recoded (Q99)			
		lower	medium	higher	
Vertrouwen_r	,00	284	57	84	425
	1,00	981	226	336	1543
Total		1265	283	420	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,028	,465
	Cramer's V	,028	,465
N of Valid Cases		1968	

B2.5: kruistabel voor vertrouwen en opleidingsniveau vader

Vertrouwen_r * educational level mother: recoded (Q100)

Count		Crosstab			Total
		educational level mother: recoded (Q100)			
		lower	medium	higher	
Vertrouwen_r	,00	316	53	56	425
	1,00	1140	197	206	1543
Total		1456	250	262	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,005	,980
	Cramer's V	,005	,980
N of Valid Cases		1968	

B2.6: kruistabel voor leeftijd en opleidingsniveau moeder

Vertrouwen_r * at age 14, parent(s) had problems making ends meet (Q104D)

Count		Crosstab				Total
		at age 14, parent(s) had problems making ends meet (Q104D)				
		yes	to some extent	a little bit	no	
Vertrouwen_r	,00	65	89	73	198	425
	1,00	183	301	281	778	1543
Total		248	390	354	976	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,049	,199
	Cramer's V	,049	,199
N of Valid Cases		1968	

B2.7: kruistabel voor vertrouwen en economische situatie gezin

Vertrouwen_r * Geslacht_r

Crosstab

Count

	Geslacht_r	Geslacht_r		Total
		,00	1,00	
Vertrouwen_r	,00	209	216	425
Vertrouwen_r	1,00	720	823	1543
Total		929	1039	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,021	,358
	Cramer's V	,021	,358
N of Valid Cases		1968	

Vertrouwen_r * educational level respondent: recoded (Q81)

Crosstab

Count

	educational level respondent: recoded (Q81)	educational level respondent: recoded (Q81)			Total
		lower	medium	higher	
Vertrouwen_r	,00	137	133	155	425
Vertrouwen_r	1,00	409	400	734	1543
Total		546	533	889	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,092	,000
	Cramer's V	,092	,000
N of Valid Cases		1968	

B2.8: kruistabel voor vertrouwen en geslacht

Vervolgens zijn kruistabellen bekeken tussen opleidingsniveau vader en de overige categorische variabelen. De syntax is als volgt:

```

CROSSTABS
  /TABLES=v262_r BY v263_r v270
  Geslacht_r v243_r
  /FORMAT=AVALUE TABLES
  /STATISTICS=PHI
  /CELLS=COUNT
  /COUNT ROUND CELL.
    
```

Hierbij hoort de volgende output:

educational level father: recoded (Q99) * educational level mother: recoded (Q100)

Crosstab

Count

	educational level mother: recoded (Q100)	educational level mother: recoded (Q100)			Total
		lower	medium	higher	
educational level father: recoded (Q99)	lower	1176	58	31	1265
educational level father: recoded (Q99)	medium	145	102	36	283
educational level father: recoded (Q99)	higher	135	90	195	420
Total		1456	250	262	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,658	,000
	Cramer's V	,465	,000
N of Valid Cases		1968	

B2.9: kruistabel voor opleidingsniveau vader en opleidingsniveau moeder

educational level father: recoded (Q99) * at age 14, parent(s) had problems making ends meet (Q104D)

Crosstab

Count

		at age 14, parent(s) had problems making ends meet (Q104D)				Total
		yes	to some extent	a little bit	no	
educational level father: recoded (Q99)	lower	205	324	252	484	1265
	medium	22	31	35	195	283
	higher	21	35	67	297	420
Total		248	390	354	976	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,318	,000
	Cramer's V	,225	,000
N of Valid Cases		1968	

B2.10: kruistabel voor opleidingsniveau vader en economische situatie gezin

educational level father: recoded (Q99) * Geslacht_r

Crosstab

Count

		Geslacht_r		Total
		,00	1,00	
educational level father: recoded (Q99)	lower	603	662	1265
	medium	137	146	283
	higher	189	231	420
Total		929	1039	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,024	,579
	Cramer's V	,024	,579
N of Valid Cases		1968	

B2.11: kruistabel voor opleidingsniveau vader en geslacht

educational level father: recoded (Q99) * educational level respondent: recoded (Q81)

Crosstab

Count

		educational level respondent: recoded (Q81)			Total
		lower	medium	higher	
educational level father: recoded (Q99)	lower	498	350	417	1265
	medium	27	103	153	283
	higher	21	80	319	420
Total		546	533	889	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,408	,000
	Cramer's V	,289	,000
N of Valid Cases		1968	

B2.12: kruistabel voor opleidingsniveau vader en opleidingsniveau respondent

Vervolgens zijn kruistabellen bekeken tussen opleidingsniveau moeder en de overige categorische variabelen. De syntax is als volgt:

```
CROSSTABS
  /TABLES=v263_r BY v270 Geslacht_r
v243_r
  /FORMAT=AVALUE TABLES
  /STATISTICS=PHI
  /CELLS=COUNT
  /COUNT ROUND CELL.
```

Hierbij hoort de volgende output:

educational level mother: recoded (Q100) * at age 14, parent(s) had problems making ends meet (Q104D)

Crosstab

Count

		at age 14, parent(s) had problems making ends meet (Q104D)				Total
		yes	to some extent	a little bit	no	
educational level mother: recoded (Q100)	lower	217	332	288	619	1456
	medium	15	30	25	180	250
	higher	16	28	41	177	262
Total		248	390	354	976	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,245	,000
	Cramer's V	,173	,000
N of Valid Cases		1968	

B2.13: kruistabel voor opleidingsniveau moeder en economische situatie gezin

educational level mother: recoded (Q100) * Geslacht_r

Crosstab

Count

		Geslacht_r		Total
		,00	1,00	
educational level mother: recoded (Q100)	lower	695	761	1456
	medium	117	133	250
	higher	117	145	262
Total		929	1039	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,021	,650
	Cramer's V	,021	,650
N of Valid Cases		1968	

B2.14: kruistabel voor opleidingsniveau moeder en geslacht

educational level mother: recoded (Q100) * educational level respondent: recoded (Q81)

Crosstab

Count

		educational level respondent: recoded (Q81)			Total
		lower	medium	higher	
educational level mother: recoded (Q100)	lower	520	392	544	1456
	medium	14	77	159	250
	higher	12	64	186	262
Total		546	533	889	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,325	,000
	Cramer's V	,229	,000
N of Valid Cases		1968	

B2.15: kruistabel voor opleidingsniveau moeder en opleidingsniveau respondent

Vervolgens zijn kruistabellen bekeken tussen de economische situatie van het gezin en de overige categorische variabele. De syntax is als volgt:

```
CROSSTABS
  /TABLES= v270 BY Geslacht_r v243_r
  /FORMAT=AVALUE TABLES
  /STATISTICS=PHI
  /CELLS=COUNT
  /COUNT ROUND CELL.
```

Hierbij hoort de volgende output:

at age 14, parent(s) had problems making ends meet (Q104D) * Geslacht_r

Crosstab

Count

		Geslacht_r		Total
		,00	1,00	
at age 14, parent(s) had problems making ends meet (Q104D)	yes	122	126	248
	to some extent	181	209	390
	a little bit	172	182	354
	no	454	522	976
Total		929	1039	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,022	,813
	Cramer's V	,022	,813
N of Valid Cases		1968	

B2.16: kruistabel voor economische situatie gezin en geslacht

at age 14, parent(s) had problems making ends meet (Q104D) * educational level respondent: recoded (Q81)

Crosstab

Count

		educational level respondent: recoded (Q81)			Total
		lower	medium	higher	
at age 14, parent(s) had problems making ends meet (Q104D)	yes	110	63	75	248
	to some extent	137	96	157	390
	a little bit	112	95	147	354
	no	187	279	510	976
Total		546	533	889	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,212	,000
	Cramer's V	,150	,000
N of Valid Cases		1968	

B2.17: kruistabel voor economische situatie gezin en opleidingsniveau respondent

Vervolgens zijn kruistabellen bekeken tussen geslacht en de overige categorische variabele. De syntax is als volgt:

```

CROSSTABS
  /TABLES=Geslacht_r BY v243_r
  /FORMAT=AVALUE TABLES
  /STATISTICS=PHI
  /CELLS=COUNT
  /COUNT ROUND CELL.

```

Geslacht_r * educational level respondent: recoded (Q81)
Crosstabulation

Count

		educational level respondent: recoded (Q81)			Total
		lower	medium	higher	
Geslacht_r	,00	239	264	426	929
	1,00	307	269	463	1039
Total		546	533	889	1968

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,045	,141
	Cramer's V	,045	,141
N of Valid Cases		1968	

B2.18: kruistabel voor geslacht en opleidingsniveau respondent

Regressieanalyses

Model 1

Allereerst zijn de controlevariabelen geslacht, leeftijd en opleidingsniveau respondent samen met de afhankelijke variabele vertrouwen in onbekenden toegevoegd aan model 1.

```
LOGISTIC REGRESSION VARIABLES Vertrouwen_r
/METHOD=ENTER Geslacht_r age OpleidingRes_M OpleidingRes_H
/PRINT=GOODFIT
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

Hierbij hoort de volgende output:

Omnibus Tests of Model Coefficients					Model Summary			
		Chi-square	df	Sig.	Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
Step 1	Step	73,434	4	,000	1	1980,138 ^a	,037	,057
	Block	73,434	4	,000				
	Model	73,434	4	,000				

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

Classification Table ^a					Hosmer and Lemeshow Test			
		Predicted		Percentage Correct	Step	Chi-square	df	Sig.
Observed	Vertrouwen_r	,00	1,00		1	13,543	8	,094
Step 1	Vertrouwen_r	,00	3	422				
		1,00	3	1540				
Overall Percentage								78,4

a. The cut value is ,500

Contingency Table for Hosmer and Lemeshow Test						
		Vertrouwen_r = ,00		Vertrouwen_r = 1,00		Total
Step 1		Observed	Expected	Observed	Expected	
1	1	70	75,799	129	123,201	199
1	2	76	58,134	121	138,866	197
1	3	40	50,728	154	143,272	194
1	4	49	46,156	145	147,844	194
1	5	41	43,155	158	155,845	199
1	6	37	38,108	156	154,892	193
1	7	33	35,758	166	163,242	199
1	8	28	30,406	165	162,594	193
1	9	26	26,201	173	172,799	199
1	10	25	20,556	176	180,444	201

Variables in the Equation							
Step 1 ^a		B	S.E.	Wald	df	Sig.	Exp(B)
	Geslacht_r	,156	,112	1,942	1	,163	1,169
	age: respondent (constructed) (Q64)	,026	,003	54,006	1	,000	1,026
	OpleidingRes_M	,364	,152	5,785	1	,016	1,440
	OpleidingRes_H	,779	,142	29,993	1	,000	2,179
	Constant	-,538	,241	5,008	1	,025	,584

a. Variable(s) entered on step 1: Geslacht_r, age: respondent (constructed) (Q64), OpleidingRes_M, OpleidingRes_H.

B2.19: output van de regressieanalyse van model 1

Model 2

Vervolgens zijn in model 2 de onafhankelijke variabelen opleidingsniveau vader en opleidingsniveau moeder aan het model toegevoegd. In totaal zijn die 4 variabelen, omdat er dummy's van zijn gemaakt. De syntax voor het realiseren van model 1 is als volgt:

```
LOGISTIC REGRESSION VARIABLES Vertrouwen_r
/METHOD=ENTER Geslacht_r age OpleidingRes_M OpleidingsRes_H
/METHOD=ENTER OpleidingVader_M OpleidingVader_H OpleidingMoeder_M
OpleidingMoeder_M
/PRINT=GOODFIT
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

Hierbij hoort de volgende output:

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	21,500	2	,000
	Block	21,500	2	,000
	Model	78,109	8	,000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1975,463 ^a	,039	,060

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

Classification Table^a

Step 1	Observed	Predicted		Percentage Correct
		Vertrouwen_r ,00	Vertrouwen_r 1,00	
Vertrouwen_r	,00	3	422	,7
	1,00	6	1537	99,6
Overall Percentage				78,3

a. The cut value is ,500

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	5,826	8	,667

Contingency Table for Hosmer and Lemeshow Test

Step 1		Vertrouwen_r = ,00		Vertrouwen_r = 1,00		Total
		Observed	Expected	Observed	Expected	
1	1	74	75,830	124	122,170	198
2	2	69	59,561	128	137,439	197
3	3	51	51,975	146	145,025	197
4	4	44	46,822	154	151,178	198
5	5	35	42,507	162	154,493	197
6	6	39	38,429	158	158,571	197
7	7	39	34,461	156	160,539	195
8	8	28	30,667	169	166,333	197
9	9	24	25,745	173	171,255	197
10	10	22	19,001	173	175,999	195

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	OpleidingVader_M	,174	,185	,889	1	,346	1,190
	OpleidingVader_H	,090	,185	,240	1	,625	1,095
	OpleidingMoeder_M	,175	,197	,790	1	,374	1,191
	OpleidingMoeder_H	,260	,214	1,470	1	,225	1,297
	sex respondent (Q63)	,157	,112	1,946	1	,163	1,169
	age: respondent (constructed) (Q64)	,029	,004	56,643	1	,000	1,029
	OpleidingRes_M	,317	,153	4,266	1	,039	1,373
	OpleidingRes_H	,686	,151	20,630	1	,000	1,985
	Constant	-,896	,311	8,307	1	,004	,408

a. Variable(s) entered on step 1: OpleidingRes_M, OpleidingRes_H.

B2.20: output van de regressieanalyse van model 2

Model 3

In model 3 zijn de onafhankelijke variabelen opleidingsniveau vader en opleidingsniveau moeder aan het model toegevoegd met als afhankelijke variabele economische situatie gezin. Er is een lineaire regressie uitgevoerd omdat de variabele economische situatie gezin bestaat uit 4 categorieën en na uitvoeren van beide analyses blijken geen grote verschillen. De syntax is als volgt:

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT v270
/METHOD=ENTER OpleidingVader_M OpleidingVader_H OpleidingMoeder_M
OpleidingMoeder_H
/METHOD=ENTER Geslacht_r age OpleidingRes_M OpleidingRes_H.
```

Hierbij hoort de volgende output:

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.302 ^a	.091	.089	1.048
2	.331 ^b	.110	.106	1.039

a. Predictors: (Constant), OpleidingMoeder_M, OpleidingVader_M, OpleidingVader_L, OpleidingMoeder_L

b. Predictors: (Constant), OpleidingMoeder_M, OpleidingVader_M, OpleidingVader_L, OpleidingMoeder_L, Geslacht_r, OpleidingRes_M, age: respondent (constructed) (Q64), OpleidingRes_L

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	221.427	4	55.357	50.364	.000 ^b
	Residual	2209.280	2010	1.099		
	Total	2430.708	2014			
2	Regression	266.241	8	33.280	30.844	.000 ^c
	Residual	2164.466	2006	1.079		
	Total	2430.708	2014			

a. Dependent Variable: at age 14, parent(s) had problems making ends meet (Q104D)

b. Predictors: (Constant), OpleidingMoeder_M, OpleidingVader_M, OpleidingVader_L, OpleidingMoeder_L

c. Predictors: (Constant), OpleidingMoeder_M, OpleidingVader_M, OpleidingVader_L, OpleidingMoeder_L, Geslacht_r, OpleidingRes_M, age: respondent (constructed) (Q64), OpleidingRes_L

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.534	.067		52.697	.000
	OpleidingVader_L	-.614	.073	-.267	-8.456	.000
	OpleidingVader_M	-.098	.085	-.031	-1.157	.247
	OpleidingMoeder_L	-.126	.085	-.050	-1.483	.138
	OpleidingMoeder_M	.133	.097	.040	1.379	.168
2	(Constant)	3.779	.096		39.461	.000
	OpleidingVader_L	-.533	.074	-.232	-7.158	.000
	OpleidingVader_M	-.093	.085	-.030	-1.103	.270
	OpleidingMoeder_L	.005	.088	.002	.059	.953
	OpleidingMoeder_M	.151	.096	.046	1.575	.115
	Geslacht_r	.019	.047	.009	.411	.681
	age: respondent (constructed) (Q64)	-.006	.002	-.099	-4.145	.000
	OpleidingRes_L	-.231	.062	-.095	-3.750	.000
	OpleidingRes_M	-.018	.058	-.007	-.311	.756

a. Dependent Variable: at age 14, parent(s) had problems making ends meet (Q104D)

B2.21: output van de regressieanalyse van model 3

Model 4

In model 4 is aan model 2 de variabele economische situatie gezin toegevoegd. Deze variabele bestaat uit drie dummyvariabelen, die in een keer zijn toegevoegd. De syntax die is gerund is als volgt:

```
LOGISTIC REGRESSION VARIABLES Vertrouwen_r
/METHOD=ENTER Geslacht_r age OpleidingRes_M OpleidingsRes_H
/METHOD=ENTER OpleidingVader_M OpleidingVader_H OpleidingMoeder_M
OpleidingMoeder_M
/METHOD=ENTER Moeite_Rondkomen_Ja Moeite_Rondkomen_Enigszins
Moeite_Rondkomen_Beetje
/PRINT=GOODFIT
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

Hierbij hoort de volgende output:

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	5.856	3	.119
	Block	5.856	3	.119
	Model	83.965	11	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1969.607 ^a	.042	.064

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Classification Table^a

	Observed	Predicted	Vertrouwen_r		Percentage Correct
			.00	1.00	
Step 1	Vertrouwen_r	.00	2	423	.5
		1.00	2	1541	99.9
Overall Percentage					78.4

a. The cut value is .500

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	5,769	8	,673

Contingency Table for Hosmer and Lemeshow Test

Step 1		Vertrouwen_r = ,00		Vertrouwen_r = 1,00		Total
		Observed	Expected	Observed	Expected	
1	1	81	76,814	116	120,186	197
2	2	62	59,909	135	137,091	197
3	3	52	52,169	145	144,831	197
4	4	44	47,192	153	149,808	197
5	5	45	43,215	155	156,785	200
6	6	29	38,623	169	159,377	198
7	7	40	34,065	157	162,935	197
8	8	31	29,919	166	167,081	197
9	9	22	25,419	176	172,581	198
10	10	19	17,675	171	172,325	190

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	OpleidingVader_L	-.024	.188	.016	1	.898	.976
	OpleidingVader_M	.102	.210	.236	1	.627	1.107
	OpleidingMoeder_L	-.266	.215	1.534	1	.216	.766
	OpleidingMoeder_M	-.097	.233	.174	1	.677	.907
	Geslacht_r	.158	.112	1.971	1	.160	1.171
	age: respondent (constructed) (Q64)	.030	.004	59.289	1	.000	1.030
	OpleidingRes_L	-.663	.152	19.098	1	.000	.515
	OpleidingRes_M	-.364	.140	6.784	1	.009	.695
	Moeite_Rondkomen_Ja	-.363	.175	4.294	1	.038	.695
	Moeite_Rondkomen_Enigszins	-.250	.155	2.622	1	.105	.779
	Moeite_Rondkomen_Beeetje	-.016	.159	.010	1	.921	.984
	Constant	.300	.226	1.761	1	.184	1.349

a. Variable(s) entered on step 1: Moeite_Rondkomen_Ja, Moeite_Rondkomen_Enigszins, Moeite_Rondkomen_Beeetje.

B2.22: output van de regressieanalyse van model 4

Bijlage 3

Modevaluatie

Deviance

Voor de modevaluatie is er gekeken naar de deviance met bijbehorende chi-kwadratescores en vrijheidsgraden. De chi-kwadratescore voor het eerste model is 1980.14 (df=4), de chi-kwadratescore voor het tweede model is 1975.46 (df=8) en de chi-kwadratescore voor het vierde model is 1969.61 (df=11). Naar mate meer variabelen aan het model worden toegevoegd, wordt de chi-kwadratescore kleiner. Dit betekent dat het model een betere voorspeller wordt van vertrouwen in onbekenden.

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1980,138 ^a	,037	,057

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

B3.1: output van de deviance (model 1)

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1975,463 ^a	,039	,060

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

B3.2: output van de deviance (model 2)

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1969.607 ^a	.042	.064

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

B3.3: output van de deviance (model 4)

Hosmer en Lemeshow Goodness-of-Fit Test

Daarnaast is er gekeken naar de Hosmer en Lemeshow Goodness-of-Fit Test voor de modellen. Met name de chi-kwadratescore voor het complete model is van belang. Dit model heeft een chi-kwadratescore van 5.769 ($p=0.673$; $df=8$). We kunnen de nulhypothese niet verwerpen en nemen aan dat het model goed bij de datapast.

Step	Chi-square	df	Sig.
1	13,543	8	,094

Step	Chi-square	df	Sig.
1	5,826	8	,667

Step	Chi-square	df	Sig.
1	5,769	8	,673

B3.4. output van de Hosmer and Lemeshow Test van model 1, model 2 en model 4

Multicollineariteit

De multicollineariteit is berekend door VIF-scores uit te rekenen. Alle variabelen zijn in het model gestopt, waarna de VIF-scores per variabele zijn genoteerd in tabel x.

B3.5: VIF-scores voor de variabelen in het complete model

Variabele	VIF
Opleidingsniveau vader	1.755
Opleidingsniveau moeder	1.728
Economische situatie	1.115
Geslacht	1.005
Leeftijd	1.239
Opleidingsniveau respondent	1.221

Er een VIF-score > 4 wijst op grote samenhang tussen een variabele en de andere variabele. Er is dan ten minste met één van de andere variabelen grote samenhang. In dit model zijn alle scores < 2 waaruit we concluderen dat er geen grote samenhang is tussen de variabelen en er lijkt dus geen sprake van multicollineariteit.

Invloedrijke punten

Om te beoordelen of er invloedrijke punten zijn die de data beïnvloeden is gekeken naar leverage scores en DFBETA-scores. De syntax is hieronder weergegeven:

```
LOGISTIC REGRESSION VARIABLES Vertrouwen_r
/METHOD=ENTER OpleidingVader_M OpleidingVader_H OpleidingMoeder_M
OpleidingMoeder_H
Moeite_Rondkomen_Ja Moeite_Rondkomen_Enigszins Moeite_Rondkomen_Beetje
Geslacht_r age OpleidingRes_M
OpleidingRes_H
/SAVE=LEVER DFBETA
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

De DFBETA-scores voor de constante zijn uitgerekend. De vuistregel is $DFBETA > 2/\sqrt{N}$ geldt als invloedrijke score. Dit geeft ons een score van 0.045 en in de dataset is te zien dat er geen scores hoger zijn dan dit getal. Om die reden zijn de DFBETA-scores niet opgenomen in de tabel.

Ook zijn er geen invloedrijke punten gevonden gebaseerd op leverage scores. Een hoge score op leverage betekent dat de score verder weg ligt van het gemiddelde en meer invloed heeft op de geschatte hellingen. Om tot resultaten te komen wordt gekeken naar extreme gevallen, maar omdat dit hier niet het geval is worden er geen scores weergegeven.