



**rijksuniversiteit
groningen**

Bachelorwerkstuk

**Biedt bestaansonzekerheid een verklaring voor
gezondheidsverschillen tussen lager en hoger opgeleiden?**

Naam Student: Melissa Sagel

Studentnummer: S4131983

Emailadres: m.sagel.1@student.rug.nl

Naam begeleider: Rita Smaniotto

Tweede lezer: Gert Stulp

Opleiding: Sociologie

Cursus: Bachelorwerkstuk

ABSTRACT

Wanneer mensen, langdurig, ziek zijn heeft dit een duurder wordende zorg en een ontwrichte samenleving tot gevolg. Er is al veel onderzoek gedaan naar het effect van opleidingsniveau op de gezondheid van mensen. Hieruit is gebleken dat lager opgeleiden over het algemeen minder gezond zijn en minder lang leven dan hoogopgeleiden. Een gebrek aan kennis van laagopgeleiden wordt meestal genoemd als reden voor dit verschil. Dit onderzoek verwacht dat het verband tussen opleidingsniveau en gezondheid deels te verklaren is door een ander mechanisme, namelijk bestaansonzekerheid. De verwachting is dat laagopgeleiden meer bestaansonzekerheid ervaren, waardoor ze minder gezond zijn dan hoogopgeleiden. Voor het onderzoeken van de onderzoeksvraag is gebruik gemaakt van al bestaande data van het LISS-panel. Dit panel neemt online vragenlijsten af bij verschillende soorten huishoudens in Nederland, om representatieve data te verzamelen over verschillende onderwerpen. Voor dit onderzoek is gebruik gemaakt van vragenlijsten over financiën, gezondheid en algemene kenmerken. De gebruikte steekproef voor dit onderzoek bestaat uit 2465 respondenten. Er is een logistische regressie, uitgevoerd om te onderzoeken of bestaansonzekerheid een verklaring kan bieden voor de slechtere gezondheid van laagopgeleiden ten opzichte van hoogopgeleiden. De resultaten waren gedeeltelijk in overeenstemming met de verwachting. Mensen met een laag opleidingsniveau ervaren inderdaad een hogere mate van bestaansonzekerheid en een slechtere gezondheid dan mensen met een hoog opleidingsniveau. Wel lijkt bestaansonzekerheid geen verklaring te bieden voor het gezondheidsverschil tussen hoog- en laagopgeleiden. Ook zijn er naar aanleiding van de resultaten implicaties opgesteld voor dit onderzoek. De belangrijkste uitkomsten van het onderzoek zijn dat lager opgeleiden zowel meer bestaansonzekerheid als een slechtere gezondheid ervaren en dat een hogere mate van bestaansonzekerheid leidt tot een minder goede gezondheid. Toch is er geen ondersteuning gevonden voor dat bestaansonzekerheid een verklaring biedt voor het gezondheidsverschil tussen hoog- en laagopgeleiden. De implicaties naar aanleiding van dit onderzoek zijn dat er gekeken moet worden hoe bestaansonzekerheid vermindert kan worden om de gezondheid te verbeteren. Ook lijkt het verschil in kennis wel een verklaring te zijn voor de opleidingskloof, dit kan een uitgangspunt zijn voor vervolgonderzoek.

INHOUDSOPGAVE

Inleiding	5
Theorie	6
Opleidingsniveau en gezondheid.....	8
Het effect van bestaansonzekerheid op het verband tussen opleidingsniveau en gezondheid	10
Controlevariabelen	11
Methode	13
Het gebruikte databestand.....	13
Onderzoeksdesign	14
Operationalisatie	14
De analyse-opzet	16
Resultaten	18
Beschrijvende statistieken.....	Error! Bookmark not defined.
Bivariate Correlaties	20
Modevaluatie.....	22
Hypothese toetsing.....	25
Conclusie	27
Discussie	28
Literatuurlijst	29
Bijlage1	
33	
Selecties uitgevoerd op data	33
Contrale samenhang van de schaalvariabele bestaansonzekerheid	33
Frequentieverdeling, descriptives en operationalisaties van de variabelen	34
<i>afhankelijke variabele gezondheid</i>	34
<i>onafhankelijke variabele opleidingsniveau</i>	37
<i>mediatie-variabele bestaansonzekerheid</i>	39
<i>controle variabele geslacht</i>	44
<i>Controle variabele geslacht</i>	45
<i>Controlevariabele chronische aandoening</i>	46

Bijlage 2	47
<i>Bivariate analyse</i>	47
<i>gemiddeldes van alle variabelen verdeeld in gezonde en minder gezonde mensen</i>	49
<i>Verdeling huishoudhoofden door geslacht</i>	51
<i>Geschonden assumpties van de lineaire regressieanalyse</i>	52
<i>logistische regressieanalyse</i>	53
<i>lineaire regressie model 3</i>	58
Bijlage 3	61
<i>Controle op outliers</i>	61
<i>Controle op multicolineariteit</i>	64

INLEIDING

Door COVID-19 is duidelijk geworden dat Nederland, nu en in de toekomst, een tekort heeft aan zorgpersoneel (NOS, 2022). Toen de druk op de zorg steeg, werd duidelijk dat er niet genoeg personeel was om dit op te vangen. Ook de lange wachtlijsten in de zorg, ook na corona, laten zien dat er een tekort is aan zorgpersoneel. Een tekort aan zorgpersoneel is niet makkelijk op te lossen, het opleiden van mensen tot zorgmedewerkers kost namelijk tijd. Wel zou er gekeken kunnen worden naar de druk op de zorg oftewel, hoe zorg je ervoor dat mensen gezonder worden? Ook wordt de zorg steeds duurder (De Groot, 2022). De steeds erger wordende vergrijzing, toenemende chronische ziekten en duurder wordende medicijnen worden genoemd als redenen voor de duurder, zelfs onbetaalbaar, wordende zorg (Hosseini, 2015). Het ziek worden van mensen is ook belastend voor de samenleving. Wanneer iemand niet normaal kan functioneren wordt de samenleving ontwricht, zowel op werkgebied als op privégebied (Pflanz & Keupp, 1977). Dit zijn goede redenen om te kijken naar wat er moet gebeuren om te zorgen dat minder mensen ziek worden en dat mensen in het algemeen gezonder worden.

Volgens het WHO is gezondheid ‘een toestand van volledig fysiek, sociaal en geestelijk welbevinden en niet van alleen het ontbreken van ziekte’ (Hoeymans, 2013). Hiermee wordt dus bedoeld dat een persoon zich op fysiek, sociaal en geestelijk vlak goed voelt en hierin geen klachten ervaart. Voor dit onderzoek zal deze definitie ook gehanteerd worden. In dit onderzoek zal het daarom gaan over de subjectieve gezondheid, dus hoe iemand zijn of haar gezondheid ervaart. Hiervoor is gekozen omdat de subjectieve gezondheid te meten is door te vragen naar hoe iemand zijn of haar gezondheid ervaart. In een ander onderzoek over gezondheid is dit concept ook gemeten doormiddel van vragenlijsten over hoe iemand zijn of haar gezondheid ervaart (Rademaker & Holland, 2021)

Om een antwoord te kunnen geven op hoe de gezondheid kan worden verbeterd is het van belang om vast te stellen wat een verklaring kan zijn voor de verschillen in gezondheid tussen mensen.

Er is al veel onderzoek gedaan naar het verschil in gezondheid tussen mensen met een hoog opleidingsniveau en mensen met een laag opleidingsniveau (Launoy, 2023; Tataw, Ekundayo & Baxter, 2022;). Met opleiding wordt de hoogst genoten opleiding van een persoon bedoeld. Onderzoek heeft uitgewezen dat hoogopgeleide mensen bijna 7 jaar langer leven dan laagopgeleiden en dat zij in het algemeen ook minder gezondheidsproblemen ervaren (Bruggink, 2009). Een veel genoemde verklaring voor het verschil in gezondheid tussen de verschillende opleidingsniveaus is het verschil in kennis. Zo zouden laagopgeleiden een minder goede gezondheid hebben omdat ze minder kennis hebben over gezondheid (Stronks et al., 2019). Doordat mensen minder kennis hebben over gezondheid, weten ze in mindere mate hoe ze ervoor moeten zorgen dat ze gezond blijven en of

worden. Denk hierbij aan kennis over voedingsmiddelen, het op tijd signaleren van lichaamssignalen en weten wanneer medische hulp gezocht moet worden.

Dit onderzoek richt zich op een ander mogelijk mechanisme dat het verschil tussen hoog- en laagopgeleiden kan verklaren. Dit mechanisme is bestaansonzekerheid. Met bestaansonzekerheid wordt bedoeld dat iemand moeite heeft met het kunnen voorzien in de basisbehoeften en hier zorgen over heeft (Rebechi & Rohde, 2022). Hierbij gaat het niet alleen over de huidige situatie, maar vooral ook over de onzekerheid of er in de toekomst financiële problemen zullen ontstaan.

Bestaansonzekerheid kan zowel subjectief als objectief worden opgevat. In dit onderzoek wordt er vooral gefocust op het subjectieve gedeelte van bestaansonzekerheid, dus op de stress en onzekerheid die worden ervaren door de financiële situatie.

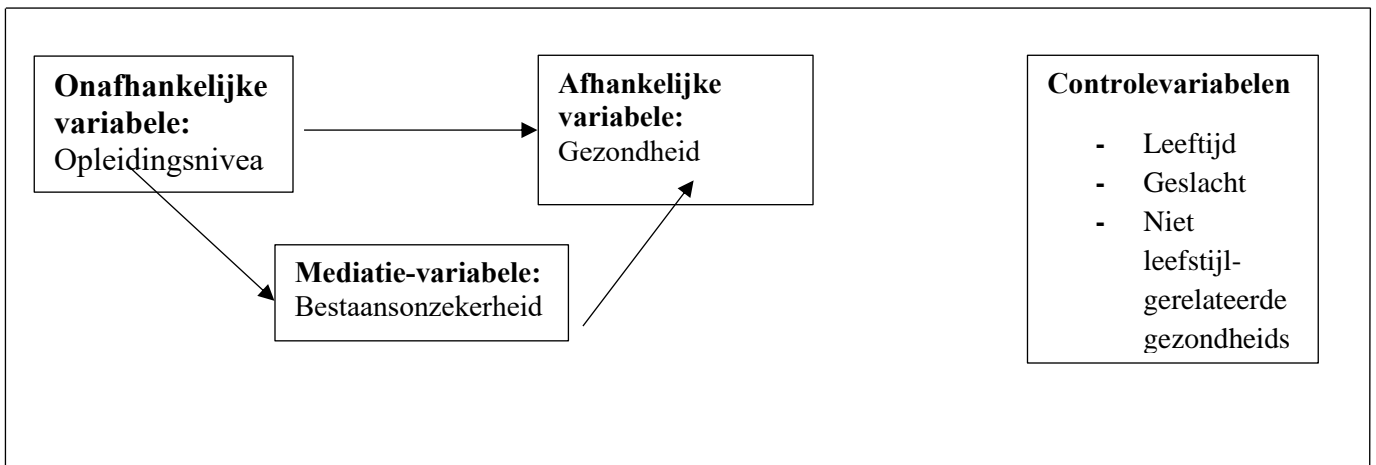
Het is relevant om te kijken of bestaansonzekerheid een verklaring zou kunnen bieden voor de verschillen in gezondheid tussen laag en hoog opgeleiden. Eerder onderzoek heeft aangetoond dat een hoger opleidingsniveau meestal een hoger inkomen genereert dan een lager opleidingsniveau (McGill, 2016). Ook vinden hoog opgeleiden vaak sneller een baan of krijgen zij een vast contract (Bussi & Pareliussen, 2017). Deze factoren maken dat de mate van bestaansonzekerheid bij laagopgeleiden hoger is dan bij hoogopgeleiden. Deze bestaansonzekerheid zou kunnen leiden, zoals eerder al genoemd, tot stress. Dit is weer ongunstig voor het ervaren van gezondheidsklachten en heeft een slechtere gezondheid tot gevolg (Kasl, 1984).

Als bestaansonzekerheid een verklaring biedt voor het verband tussen gezondheid en opleidingsniveau is dit een belangrijke toevoeging aan onze kennis over de verschillen in gezondheid tussen hoog- en laagopgeleiden. Op dit moment is de dominante interventie het voorlichting geven over een gezonde leefstijl (Whitehead & Dahlgren, 1991). Omdat gebrek aan kennis wordt gezien als reden voor gezondheidsverschillen tussen hoog- en laagopgeleiden, wordt voorlichting gezien als de beste interventie. Als de gezondheidsverschillen tussen hoog- en laagopgeleiden het gevolg zijn van bestaansonzekerheidsverschillen, betekent dit dat die interventie niet effectief is. In plaats daarvan zouden interventies zich moeten richten op manieren om de bestaansonzekerheid van laagopgeleiden te verkleinen. Er zou dan kunnen worden ingezet op programma's om bestaansonzekerheid te verkleinen, denk hierbij bijvoorbeeld aan het creëren van meer vaste banen voor laagopgeleiden. Dit onderzoek kan daarom relevante informatie toevoegen over de rol die bestaansonzekerheid speelt in de gezondheidsverschillen tussen hoog-en laagopgeleiden.

De probleemstelling die voortvloeit uit de eerdergenoemde alinea's van dit hoofdstuk is: kan bestaansonzekerheid een verklaring bieden voor de lagere gezondheid van laagopgeleiden ten opzichte van hoogopgeleiden? Dit onderzoek gaat proberen een antwoord te geven op deze

probleemstelling, door gebruik te maken van data van het LISS-panel. Dit is een langlopend onderzoek waarin respondenten gevolgd worden gedurende een lange tijd (LISS panel, 2022).

THEORIE



Figuur 1, conceptueel model

Zoals staat beschreven in de inleiding, is de onderzoeksvraag van dit onderzoek: ‘kan bestaansonzekerheid een verklaring bieden voor het verband tussen opleidingsniveau en gezondheid?’. Figuur 1 is een schematische weergave van het bijpassende onderzoeksmodel. Zoals te zien is in figuur 1 worden er ook drie controlevariabelen, leeftijd, geslacht en chronische aandoening, meegenomen in het model. In deze paragraaf zal nader in worden gegaan op de relatie tussen opleidingsniveau en gezondheid, de invloed van bestaansonzekerheid hierop en de controlevariabelen.

OPLEIDINGSNIVEAU EN GEZONDHEID

Zoals in de inleiding al werd genoemd, is er reden om aan te nemen dat een hoger opleidingsniveau leidt tot een betere gezondheid. In deze paragraaf zal er eerst ingegaan worden op een aantal veelvoorkomende verklaringen in de literatuur.

Een verschil in kennis tussen hoog- en laagopgeleiden kan een verklaring bieden voor het gezondheidsverschil tussen hoog- en laagopgeleiden (Stronks et al., 2019). Doordat lager opgeleiden minder kennis hebben over gezondheid, weten ze in mindere mate hoe ze ervoor moeten zorgen dat ze gezond blijven en of worden. Denk hierbij aan kennis over voedingsmiddelen, het op tijd signaleren van lichaamssignalen en weten wanneer medische hulp gezocht moet worden.

Een eerste verklaring is de “Fundamental Cause” theorie (Link & Phelan, 2002). Deze theorie stelt dat mensen met een hogere sociaaleconomische status meer kunnen profiteren van gezondheid verbeterende kennis en nieuwe medische innovaties vanwege een grotere toegang tot bepaalde

bronnen. Deze bronnen zijn onder andere kennis, geld en macht. De theorie stelt dat mensen met een hogere sociaaleconomische status toegang hebben tot de beste dokters, weten over en vragen naar gunstige gezondheidsbehandelingen en vaker mensen om zich heen hebben die een gezonde levensstijl kunnen promoten en ondersteunen. Mensen met een hogere sociaaleconomische status hebben dus door hun sociale omgeving en middelen, meer kennis over gezondheid (O'Donnell, 2020). Sociaaleconomische status wordt bepaald op basis van inkomen, positie op de arbeidsmarkt en opleidingsniveau (Jiang, 2023). Hoger opgeleiden hebben dan dus een hogere sociaaleconomische status dan lager opgeleiden en meer toegang tot bronnen zoals kennis, geld en macht.

Een andere verklaring is dat hoogopgeleiden over het algemeen meer lezen dan laagopgeleiden, zowel voor school als in hun vrije tijd (Gökbulut & Yeniasir, 2018). Hierdoor hebben ze meer kennis over verschillende onderwerpen, zoals kennis over gezondheid. Ook hebben lager opgeleiden minder kennis over gezondheid dan hoger opgeleiden, omdat ze informatie over gezondheid minder goed begrijpen en in mindere mate weten hoe ze die kennis moeten toepassen (Haas & Fosse, 2008). Hoogopgeleiden hebben meer ervaring met het verwerken van veel informatie doordat ze, bijvoorbeeld voor hun opleiding, meer oefenen met het verwerken van informatie dan laagopgeleiden (Gökbulut & Yeniasir, 2018). Zo is er bijvoorbeeld veel informatie te vinden over diëten en gezond eten, maar laagopgeleiden vinden het moeilijk deze informatie te filteren en te begrijpen hoe ze het nou moeten aanpakken. Een verklaring hiervoor is dat laagopgeleiden veel boeken en rapporten moeten lezen en daar de belangrijkste informatie uit moeten kunnen filteren. Laagopgeleiden doen meestal opleidingen die ze voorbereid op een fysiek beroep. Dit houdt in dat zij veel minder hoeven te lezen en veel meer in de praktijk moeten doen.

De verschillende soorten werk die mensen met hogere en lagere opleidingsniveaus uitvoeren kan ook een verklaring bieden voor de verschillen in gezondheid tussen hoog- en laagopgeleiden. Zo hebben hoger opgeleiden vaker een veiligere baan dan laagopgeleiden (Leigh, 1983). Met veiliger wordt hier voornamelijk bedoeld, veiliger voor hun gezondheid. Zo hebben hoger opgeleiden vaker kantoorbanen, terwijl lager opgeleiden fysiekere beroepen hebben, bijvoorbeeld stratenlegger, asfaltwerkers of havenwerkers. Deze fysieke beroepen zijn schadelijker voor de lange termijn, voornamelijk fysieke, gezondheid dan een kantoorbaan. Zo hebben stratenmakers bijvoorbeeld vaak last van versleten knieën als ze wat ouder worden. Er moet wel enige nuance worden aangebracht aan dit argument. Zoals eerder benoemd hebben hoger opgeleiden vaker een kantoorbaan dan hoger opgeleiden. Bij een kantoorbaan moeten mensen wel meer stil zitten, wat ook weer klachten kan veroorzaken. Wel zijn dit vaker klachten die geen langdurige klachten veroorzaken, denk hierbij aan nekklachten die een fysio kan verhelpen (Leigh, 1983). Klachten als een versleten knie, die laagopgeleiden kunnen oplopen door hun werk, zijn niet meer te verhelpen.

Hypothese 1: Hoger opgeleiden hebben een betere gezondheid dan lager opgeleiden

HET EFFECT VAN BESTAANSONZEKERHEID OP HET VERBAND TUSSEN OPLEIDINGSNIVEAU EN GEZONDHEID

Dat hoger opgeleiden gezonder zijn dan lager opgeleiden kent dus meerdere verklaringen, zoals in de vorige paragraaf is uitgelegd. Toch is er reden om aan te nemen dat het gezondheidsverschil tussen hoger en lager opgeleiden een andere verklaring kent. Deze verklaring is bestaansonzekerheid. Zoals in de inleiding al is beschreven, wordt met bestaansonzekerheid bedoeld dat iemand moeite heeft met het kunnen voorzien in de basisbehoeften, wat zorgen oplevert, vooral met betrekking tot de toekomst (Rebecchi & Rohde, 2022). In deze paragraaf is het effect van bestaansonzekerheid op de verschillen in gezondheid tussen hoog- en laagopgeleiden uitgewerkt.

Het verschil in onzekerheid op de arbeidsmarkt tussen hoog- en laagopgeleiden kan een verklaring bieden voor de hogere mate van bestaansonzekerheid bij lager opgeleiden. Kapitalisme en globalisering hebben ervoor gezorgd dat er meer flexibel-, tijdelijk- en parttimewerk is (Standing, 2011). Dit komt doordat winst steeds belangrijker is geworden en er ook veel werk naar het buitenland is verplaatst. Deze manier van werken, -met veel flexibel-, tijdelijk, en parttimewerk -zorgt voor een hogere mate van onzekerheid. Dit komt doordat er geen vast inkomen is of geen vaste aanstelling, waardoor er niet 'genoeg' verdiend wordt. Onderzoek heeft aangetoond dat hoger opgeleiden een hogere baankans hebben en een grotere kans op een vaste baan (Bussi & Pareliussen, 2017). Wanneer hoger opgeleiden een tijdelijke positie hebben vinden zij snel een andere baan en maken ze zich minder zorgen over het vinden van een nieuwe baan. Lager opgeleiden hebben deze zorgen vaak wel, voor hen is het ook moeilijker om een vaste of fulltimebaan te vinden. Lager opgeleiden leven dus meer in bestaansonzekerheid omdat ze flexibel-, tijdelijk- of parttimewerk hebben. Ook omdat de banen die ze hebben vaak al niet heel goed betaald zijn, is het extra belastend als ze niet veel kunnen werken of maar tijdelijk kunnen werken (McGill, 2016).

Daarnaast maakt de huidige tijd, waarin de prijzen fors gestegen zijn ook dat laagopgeleiden meer bestaansonzekerheid ervaren. Door de oorlog in Oekraïne zijn de prijzen van onder andere boodschappen, benzine en energie gestegen (Ministerie van Algemene Zaken, 2022). Dit leidt tot een stijgende inflatie en een dalende koopkracht. Vooral mensen met een minimaal inkomen, wat voornamelijk laagopgeleiden zijn, lijken hier het slachtoffer van te zijn. Laagopgeleiden met een laag inkomen hebben over het algemeen al moeite met rondkomen, vooral als ze een gezin hebben (Ring, 1985). Hun salaris ging ook voor de inflatie vaak al op aan vaste lasten en noodzakelijk bestedingen, zoals bijvoorbeeld boodschappen. Dit houdt in dat ze maandelijks sowieso al niet veel geld overhielden. Nu de prijzen zijn gestegen, groeit daarmee ook de onzekerheid bij laagopgeleiden of ze wel rond kunnen komen. Onvoorziene uitgaven zoals zorgkosten of een ketel die stuk gaat, zorgen bij laagopgeleiden dan voor nog meer stress en onzekerheid en dit leidt dan dus tot meer bestaansonzekerheid.

Door de zorgen die laagopgeleiden hebben over geld en onzekerheid over of ze iets wel of niet kunnen betalen, hebben ze vaak ook meer stress dan hoogopgeleiden (Zhou et al., 2017). Stress heeft een negatieve invloed op de gezondheid. Onderzoek heeft aangetoond dat stress kan leiden tot zowel fysieke als mentale klachten (Kasl, 1984). Zo kan stress er bijvoorbeeld voor zorgen dat iemands immuunsysteem wordt aangetast, waardoor zelfs ziektes het gevolg kunnen zijn van het langdurig ervaren van stress. Ook kan stress ervoor zorgen dat iemand slecht slaapt, waardoor diegene uiteindelijk moeite kan ervaren met functioneren in het dagelijkse leven en mentale problemen kan krijgen (You-Wei Yan et al., 2018). Daarom zal iemand die veel bestaansonzekerheid ervaart dus een relatief slechtere gezondheid hebben dan iemand die weinig bestaansonzekerheid ervaart.

Ook gebruiken mensen die meer bestaansonzekerheid ervaren vaker verdovende middelen, zoals alcohol en drugs, dit doen ze om even niet na te hoeven denken over hun zorgen (Kerssemakers et al., 2015). De lange termijn gevolgen van deze slechte gewoontes kunnen uiteenlopen van het krijgen van chronische ziekten tot zelfs het eerder overlijden. Zo kan langdurig roken uiteindelijk longkanker tot gevolg hebben en is intensief alcoholgebruik slecht voor de lever (Stronks et al., 2019). Ook op korte termijn ervaren mensen een minder goede gezondheid als ze verdovende middelen nemen, omdat ze vaker last hebben oververmoeidheid, het verliezen van gewicht en hoofdpijn. Iemand die door het ervaren van bestaansonzekerheid verdovende middelen neemt, zal daarom ook een slechtere gezondheid ervaren.

H2: De invloed van opleidingsniveau op gezondheid wordt (gedeeltelijk) verklaard doordat mensen met een lage opleiding meer bestaansonzekerheid ervaren

CONTROLEVARIABLEN

Verschillende factoren kunnen invloed uitoefenen op de gezondheid van mensen, ongeacht hun opleidingsniveau. Door voor deze variabelen te controleren kan er met meer zekerheid iets worden gezegd over het causale verband tussen opleidingsniveau en gezondheid en kunnen andere invloeden uitgesloten worden. Deze controlevariabelen zijn leeftijd, geslacht en niet leefstijl-gerelateerde gezondheidsproblemen.

De eerste controlevariabele is leeftijd. Mensen die ouder zijn ervaren een slechtere gezondheid dan mensen die jonger zijn (Lagaay, 1999). Ouderen ervaren vaker lichamelijke klachten dan jongeren. Deels dragen leefstijl factoren op jongere leeftijd hier ook aan bij, maar er zijn veel oude mensen die gezond leven en alsnog met deze ziektes te maken krijgen. Zo kan iemand van 85 met een hoog opleidingsniveau toch een slechte gezondheid ervaren. Om de invloed van leeftijd weg te filteren controleer ik voor leeftijd.

De tweede factor is geslacht. Onderzoek heeft aangetoond dat mannen zich over het algemeen gezonder voelen dan vrouwen, maar dat mannen wel minder oud worden dan vrouwen (Kronenfeld, 1999). Dit verschil is te verklaren doordat vrouwen eerder naar de huisarts gaan dan mannen en sneller gealarmeerd worden door lichamelijke klachten. Dit is wel opvallend aangezien mannen wel sneller als “aansteller” worden gezien als ze ziek zijn, omdat ze dan vaak meer zeuren dan vrouwen. Mannen reageren dan vaak “aanstellerig” als ze ziek zijn, maar zijn minder snel echt gealarmeerd dan vrouwen. Als er meer vrouwen in de steekproef zouden voorkomen, zou het heel goed kunnen dat, ongeacht het opleidingsniveau, de algemene gezondheid toch slechter zou zijn dan wanneer er veel mannen in de steekproef zouden voorkomen. Om de invloed van geslacht weg te filteren controleer ik voor geslacht.

De laatste factor is niet leefstijl-gerelateerde gezondheidsproblemen. Wanneer iemand niet leefstijl-gerelateerde gezondheidsproblemen ervaart die niks met de gevolgen van een opleidingsniveau te maken heeft, kan dit ook invloed hebben op de gezondheid. Zo kan iemand een auto-ongeluk hebben gehad of al een afwijking hebben sinds zijn of haar geboorte. Deze gezondheidsproblemen hebben dan niks te maken met de gevolgen van een opleidingsniveau, maar hebben wel invloed op de gezondheid. Om de invloed van niet leefstijl-gerelateerde gezondheidsproblemen weg te filteren controleer ik voor niet-leefstijl gerelateerde gezondheidsproblemen.

METHODE

In dit hoofdstuk wordt een beschrijving weergegeven van hoe dit onderzoek is uitgevoerd. Om dit op een duidelijke en transparante manier te doen, wordt de beschrijving van het databestand, het onderzoeksdesign, de operationalisaties en de analyse-opzet gegeven.

HET GEBRUIKTE DATABESTAND

Voor het verzamelen van data voor dit onderzoek is er gebruikt gemaakt van het LISS-panel. Dit is een langlopend onderzoek waarin respondenten gevolgd worden gedurende een lange tijd. Het doel van het LISS-panel is om data van de Nederlandse bevolking te verzamelen over allerlei onderwerpen, voor maatschappelijk, beleidsrelevant en wetenschappelijk onderzoek (*LISS panel*, 2022). Het LISS-panel verzamelt deze data om inzicht te krijgen in verschillende uiteenlopende onderwerpen. Het LISS-panel stelt vragen over verschillende onderwerpen, zoals gezondheid, persoonlijkheid, inkomen en school gerelateerde vragen. De populatie bestaat uit Nederlandse huishoudens. Het LISS-panel bestaat uit 4800 huishoudens. Deze huishoudens worden geselecteerd door het CBS. Het LISS-panel heeft van het CBS 10.000 adressen gekregen, waarvan er 48% uiteindelijk lid werd. Voor dit onderzoek is er gebruikt gemaakt van de modules wave 6 en wave 7 van het LISS-panel, die afgenomen zijn in 2013-2014.

Het totale aantal respondenten in de voor dit onderzoek gebruikte data is 10261. De algemene respons bij de gebruikte dataset is geen vast aantal, omdat er verschillende vragenlijsten zijn gebruikt. De vragenlijst Health LISS Core Study Health Wave 7, die wordt gebruikt voor gezondheid en niet leefstijl-gerelateerde gezondheidsproblemen, is in totaal naar 6217 mensen toegestuurd. Sowieso is deze vragenlijst alleen naar panelleden gestuurd die 16 jaar of ouder waren. De vragenlijst had een response van 5379 en een non-response van 838, dit is dus een relatief lage non-response. De reden voor de non-respons was niet vermeld, maar dit kan te maken hebben met het te confronterend vinden om te vertellen over slechte gezondheid. Iets minder dan de helft van de respondenten uit de gebruikte dataset geeft dus geen antwoord op de vragen over gezondheid of hebben de vragenlijst nooit toegestuurd gekregen.

De Income LISS Core Study Wave 6, gebruikt voor bestaansonzekerheid, is naar 6395 mensen toegestuurd. Ook deze vragenlijst is alleen naar panelleden gestuurd die 16 jaar of ouder waren. De vragenlijst heeft een response van 5015 en een non-response van 1380. Dit is een redelijk hoge non-respons, dit zou te maken kunnen hebben met dat vragen over inkomen gevoelig kunnen liggen. Schaamte kan er hier daarom voor hebben gezorgd dat mensen de vragenlijst niet hebben willen invullen. Meer dan de helft van de respondenten uit de gebruikte dataset geeft dus geen antwoord op de vragen over inkomen of hebben de vragenlijst nooit toegestuurd gekregen.

Van de vragenlijst Achtergrondvariabelen LISS-Panel, die de achtergrondvariabelen

leeftijd en geslacht weergeeft, is bekend dat de contactpersoon van een huishouden elke maand een vragenlijst over de belangrijkste algemene kenmerken ontvangt. Elk huishouden heeft 1 contactpersoon, wat inhoudt dat er 4800 mensen de vragenlijst hebben ingevuld. Dit is minder dan de helft van het totale aantal respondenten in de gebruikte dataset.

Het totale aantal respondenten in de voor dit onderzoek gebruikte data is 10261. Van de 10261 respondenten zijn er uiteindelijk 2465 opgenomen in de analyse die is gedaan voor dit onderzoek. In eerste instantie zijn er, zoals in de vorige alinea uitgelegd, grote aantallen respondenten geweest die bepaalde vragenlijsten niet hebben ontvangen of niet hebben ingevuld. Dit houdt in dat die respondenten sowieso niet in de steekproef terecht zijn gekomen. Ook zijn er nog meerdere selecties uitgevoerd. Allereerst is er geselecteerd op huishoudhoofden. In het LISS-panel doen vaak meerdere mensen uit hetzelfde huishouden mee, hierdoor is er geen sprake van onafhankelijke cases. Door te selecteren op huishoudhoofd kan dit probleem verholpen worden. Nadat er geselecteerd was voor huishoudhoofden 4327 respondenten over in de steekproef. Ook bestaat de data die wordt gebruikt voor de analyse uit respondenten die antwoord hebben gegeven op alle vragen. Er zijn dus respondenten die niet meegenomen zijn in de analyse, omdat ze niet op alle vragen antwoord hebben gegeven. Een reden hiervoor kan zijn dat meerdere vragen die meegenomen zijn in de analyse gaan over inkomen en gezondheid, deze onderwerpen kunnen gezien worden als gevoelige onderwerpen. Ook is het voor de schaalvariabele bestaansonzekerheid nodig dat er op minimaal 3 vragen, die samen de schaalvariabele vormen, antwoord wordt gegeven. Als iemand maar op één of twee vragen antwoord heeft gegeven wordt diegene al niet meegenomen in de analyse. Het uiteindelijke aantal respondenten, 2465, meegenomen in de analyse is een groot aantal. Deze groep is groot genoeg om uitspraken over te kunnen doen.

ONDERZOEKSDESIGN

Zoals in de vorige paragraaf al werd uitgelegd is er voor dit onderzoek gebruik gemaakt van het LISS-panel. Dit panel maakt gebruik van vragenlijstonderzoek om hun data te verzamelen. Het gaat hierbij om gestructureerde vragenlijsten die aan de respondenten worden voorgelegd. De vragenlijsten bestaan uit vragen over één bepaald onderwerp, bijvoorbeeld inkomen. De vragen binnen die lijst kunnen variëren van vragen over hoeveel iemand verdient tot hoe inkomen invloed heeft op het dagelijks leven van de respondent, dus of iemand goed kan rondkomen. De respondenten krijgen maandelijks vragenlijsten toegestuurd via het internet (*LISS panel, 2022*). Hiervoor ontvangt de respondent een welkomst-email met een link naar een online omgeving waarin ze maandelijks een vragenlijst vinden over verschillende onderwerpen. Maandelijks ontvangt de respondent dan een mail waarin aangegeven staat dat er een vragenlijst online staat. Het invullen van de vragenstellingen duurt niet meer dan een half uur en is anoniem.

OPERATIONALISATIE

De afhankelijke variabele gezondheid is gemeten door één vraag. De vraag is: “hoe zou over het algemeen uw gezondheid noemen?”. De antwoordmogelijkheden op deze vraag gaan van 1 tot en met 5. Waarbij 1 staat voor “slecht”, 2 voor “matig”, 3 voor “goed”, 4 voor “zeer goed” en 5 voor “uitstekend”. Omdat er uiteindelijk is gekozen voor een logistische regressie (uitleg te vinden in de analyse opzet) is van gezondheid een dummyvariabele gemaakt. Dit houdt in dat de antwoordcategorieën zijn samengevoegd tot 2 groepen. Een score van 0 staat nu voor antwoord categorieën 1 “slecht” en 2 “matig”. En een score van 1 staat nu voor antwoordcategorieën 3 “goed”, 4 “zeer goed” en 5 “uitstekend”. Oftewel een score van 0 op gezondheid staat nu voor een minder goede gezondheid en een score van 1 voor een betere gezondheid. Voor de verdeling van deze twee categorieën is gekozen mede door de verdeling de vraag, deze is te vinden in bijlage 1.

Opleidingsniveau is gemeten door één vraag. Deze vraag is: “wat is uw hoogst genoten opleiding?”. De scores van deze variabele gaan van 1 tot en met 6, waarbij 1 staat voor “basisschool”, 2 voor “vmbo”, 3 voor “havo/vwo”, 4 voor “mbo”, 5 voor “hbo” en 6 voor “wo”. Het gaat dus om een ordinale variabele.

Er worden meerdere vragen gebruikt om bestaansonzekerheid te meten. De eerste vraag is: “Ieder huishouden heeft maandelijks vaste woonlasten. Niet iedereen kan die uitgaven even makkelijk opbrengen. Hoe zwaar vallen de woonlasten voor uw huishouden?”. De score 1 staat voor “helemaal niet zwaar” en de score 7 voor “heel zwaar”. De tweede vraag is: “Kunt u op een schaal van 0 tot 10 aangeven hoe moeilijk of makkelijk u kunt rondkomen van uw inkomen?”. De score mogelijkheden gaan van 1 is “heel zwaar” naar 10 is “makkelijk”. De derde vraag is: “Hoe zou u op dit moment de financiële situatie van uw huishouden beschrijven?”. De antwoordmogelijkheden zijn 1 tot en met 5. Waarbij 1 staat voor “er worden schulden gemaakt”, 2 voor “de spaarmiddelen worden enigszins aangesproken”, 3 voor “er wordt precies rondgekomen”, 4 voor “er wordt een beetje geld overgehouden” en 5 voor “er wordt precies rondgekomen. De vierde vraag, in dit geval stelling, is: “Kijk naar de komende 12 maanden. Denkt u dan dat de uitgaven van uw huishouden: (hierop volgen 5 keuzemogelijkheden)”. Een score van 1 staat voor “veel hoger zullen zijn dan de inkomsten”, 2 voor “hoger zullen zijn dan de inkomsten”, 3 voor “ongeveer gelijk zullen zijn aan de inkomsten”, 4 voor “lager zullen zijn dan de inkomsten” en 5 voor “veel lager zullen zijn dan de inkomsten”.

Voor deze vragen is gekozen om een zo duidelijk mogelijk beeld weer te geven van bestaansonzekerheid. Zoals eerder al is benoemd, wordt met bestaansonzekerheid bedoeld dat iemand moeite heeft met het kunnen voorzien in de basisbehoeften en hier zorgen over ervaart (Rebechi & Rohde, 2022). De hierboven genoemde vragen gaan dan ook over of iemand moeilijk of makkelijk rond kan komen en hoe diegene in de toekomst verwacht rond te kunnen komen. De vraag over hoe zwaar de maandelijkse lasten iemand vallen, laat zien of iemand hier ook zorgen van ervaart.

Van de bovengenoemde variabelen is een schaalvariabele gemaakt. Hiervoor zijn de volgende

vragen omgedraaid: “kunt u op een schaal van 0 tot 10 aangeven hoe moeilijk of makkelijk u kunt rondkomen van uw inkomen?”, “hoe zou u op dit moment de financiële situatie van uw huishouden beschrijven?” en “kijk naar de komende 12 maanden. Denkt u dan dat de uitgaven van uw huishouden zijn: (hierop volgen 5 keuzemogelijkheden).”. Hiervoor is gekozen omdat alle variabelen dan dezelfde kant op zijn gecodeerd. Hoe hoger iemand scoort op bestaansonzekerheid, hoe meer bestaansonzekerheid die persoon dan ook ervaart. Ook zijn alle variabelen zo bewerkt dat ze allemaal uit 5 categorieën bestaan, dit is beter wanneer er een schaalvariabele van gemaakt wordt. De cronbach's alpha van deze schaalvariabele is 0,773. Dit houdt in dat de interne consistentie goed is en dat de afzonderlijke items goed bij elkaar passen.

De vraag die de leeftijd van de respondent meet is de vraag: “wat is uw leeftijd?”. Het antwoord op deze vraag kan, in principe, variëren van 1 tot en met 100. Het antwoord is de werkelijke leeftijd van de respondent. Het gaat hier dus om een continue variabele.

De controlevariabele niet leefstijl-gerelateerde gezondheidsproblemen is gemeten door één vraag. Deze vraag is: “Hebt u last van één of andere langdurige ziekte, aandoening, handicap of lijdt u aan de gevolgen van een ongeluk?” Waarbij 1 een “ja” is en 2 een “nee”. Het gaat hierbij om een ordinale variabele.

Het controlevariabele geslacht is gemeten door één vraag. Deze vraag is: “wat is uw geslacht?”. De bijbehorende scores zijn 1=“man” en 2=“vrouw”. Het gaat hierbij om een nominale variabele.

DE ANALYSE-OPZET

In eerste instantie is een lineaire regressieanalyse uitgevoerd, maar er werden te veel assumpties geschonden voor deze analyse (zie bijlage 3 voor de geschonden assumpties uitleg). Daarom is een logistische regressieanalyse uitgevoerd. In totaal zijn er 4 modellen getoetst. In modellen 1, 2 en 4 is gezondheid de afhankelijke variabele, deze modellen zijn geschat door logistische regressieanalyse. Model 3 is het mediatiemodel, in dit model is bestaansonzekerheid de afhankelijke variabele. Voor dit model is wel een lineaire regressieanalyse uitgevoerd.

In het eerste model is gezondheid geschat op basis van het onafhankelijke variabele opleidingsniveau. Dit model geeft het effect van opleidingsniveau op gezondheid weer, zonder invloed van de controlevariabelen. Wanneer dit model dan wordt vergeleken met het model waar de controlevariabelen aan toe zijn gevoegd, is het effect van de controlevariabelen ook goed zichtbaar.

In het tweede model zijn de controle variabelen toegevoegd aan het eerste model. Gezondheid wordt in dit model dus geschat op basis van de onafhankelijke variabele opleidingsniveau en de controle variabelen leeftijd, geslacht en niet leefstijl-gerelateerde gezondheidsproblemen. Dit model toetst dus hypothese 1: “hoger opgeleiden hebben een betere gezondheid dan lager opgeleiden”.

Aan de hand van onder andere modellen 3 en 4 kan gekeken worden of er sprake is van een

mediatie-effect. Hiervoor zijn drie eisen waaraan voldaan moet worden. Allereerst moet er sprake zijn van een effect van opleidingsniveau op bestaansonzekerheid, lager opgeleiden moeten meer bestaansonzekerheid ervaren. Dit effect kan bekeken worden in model 3. In dit model wordt getoetst of opleidingsniveau een effect heeft op bestaansonzekerheid. Dit model schat bestaansonzekerheid dus door de onafhankelijke variabele opleidingsniveaus en de controle variabelen leeftijd, geslacht en niet leefstijl-gerelateerde gezondheidsproblemen.

Vervolgens moet bestaansonzekerheid een effect hebben op gezondheid, oftewel mensen die meer bestaansonzekerheid ervaren zijn minder gezond dan mensen die in mindere mate bestaansonzekerheid ervaren. Dit effect wordt getoetst in model 4. In dit model is de mediatie variabele, bestaansonzekerheid toegevoegd aan model 2. De kans op gezondheid wordt in dit model dus geschat op basis van het onafhankelijke variabele opleidingsniveau, de mediatie variabele bestaansonzekerheid en de controle variabelen leeftijd, geslacht en niet leefstijl-gerelateerde gezondheidsproblemen

Als laatste moet het effect van opleidingsniveau op gezondheid afnemen of zelfs helemaal verdwijnen als bestaansonzekerheid is toegevoegd in model 4. Dit kan bekeken worden door model 4 met model 2 te vergelijken. Als er in model 2 een verband bestaat tussen gezondheid en opleidingsniveau, die afneemt of zelfs helemaal verdwijnt wanneer bestaansonzekerheid wordt toegevoegd in model 4, kan een mediatie-effect vastgesteld worden.

RESULTATEN

De resultaten uit het onderzoek staan centraal in dit hoofdstuk. Allereerst zullen de beschrijvende statistieken en de bivariate correlaties worden gegeven, om inzicht te geven in de verdeling van de variabelen en de samenhang tussen de variabelen. Vervolgens volgt de modevaluatie, die de kwaliteit van de getoetste modellen weergeeft. Ook zal er in deze subparagraaf een bondige beschrijving worden gegeven van de assumpties, multicollineariteit en de uitbijters. Als laatste zal er worden ingegaan op de hypothesetoetsing en de hierbij horende resultaten.

BESCHRIJVENDE STATISTIEKEN

Tabel 1: *Beschrijving van de in de analyse opgenomen variabelen: gemiddelde (standaarddeviatie), minimum- en maximumwaarde, eerste- en derde kwartiel en mediaan; N=2465*

<i>Variabele</i>	<i>Gemiddelde (Standaarddeviatie)</i>	<i>Minimum</i>	<i>Maximum</i>	<i>1^e Kwartiel</i>	<i>Mediaan</i>	<i>3^e Kwartiel</i>
Gezondheid						
<i>Goed</i>	80,5%					
<i>Minder Goed</i>	19,5%					
Opleidingsniveau	3.68(1.53)	1	6	2	4	5
Geslacht						
<i>Man</i>	69,5%					
<i>Vrouw</i>	30,5%					
Leeftijd	55.84(15,10)	18			56	66
Niet leefstijl gerelateerde gezondheidsproblemen						
	35,4%		92	44		
<i>Ja</i>						
	64,6%					
<i>Nee</i>						
Bestaansonzekerheid	2,62(0,83)	1			2	2.50
				5		3.25

In tabel 1 zijn de beschrijvende statistieken te zien. In deze sub-paragraaf worden de belangrijkste en opvallendste resultaten besproken. Allereerst valt het op dat er veel meer mensen met een goede gezondheid (80,5%) in de steekproef zijn opgenomen dan mensen met een minder goede gezondheid

(19,5%). Dit verschil zou verklaard kunnen worden doordat het om een dummy variabele gaat en er maar 2 keuzes zijn. In bijlage 1 is te zien dat de oorspronkelijke variabele gezondheid wel redelijk normaal verdeeld is. Uiteindelijk is er toch gekozen voor een logistische regressieanalyse, waardoor de variabele gezondheid een dummyvariabele is geworden. Het grote verschil in aantallen tussen mensen die minder gezond en gezond zijn zal wel meegenomen moeten worden wanneer de betrouwbaarheid van de resultaten wordt besproken.

Ten tweede valt het op dat er meer mensen in de steekproef zijn opgenomen die geen gezondheidsproblemen hebben die niet leefstijl gerelateerd zijn (64,6%) dan mensen die wel niet leefstijl gerelateerde gezondheidsproblemen hebben (35,4%). Dit verschil is niet zo groot dat het gelijk een probleem is. Ook is de variabele niet leefstijl gerelateerde gezondheidsproblemen een controle variabele en geen variabele die direct betrekking heeft tot het mediatieverband. Wel is het iets om in gedachten te houden bij het interpreteren van de resultaten.

Ook valt het op dat er meer mannen (69,5%) dan vrouwen (30,5%) zijn opgenomen in de steekproef. Dit verschil zou verklaard kunnen worden doordat er geselecteerd is op huishoudhoofd, de meeste huishoudhoofden zijn namelijk mannen. Van de 2465 huishoudhoofden zijn er namelijk 1714 een man (zie bijlage 2). Ook voor deze variabele geldt dat de variabele geen directe betrekking heeft tot het mediatieverband, maar dat dit wel iets is om mee te nemen bij het interpreteren van de resultaten.

Verder is het ook belangrijk om goed naar de variabelen opleidingsniveau en bestaansonzekerheid te kijken, omdat deze variabelen samen met gezondheid het mediatie-verband vormt die de basis is voor de hoofdvraag. Opleidingsniveau lijkt redelijk normaal verdeeld te zijn. Het heeft een gemiddelde van 3,68; terwijl de minimale score 1 is en de maximale score 6. Dit houdt in dat er dus ongeveer evenveel hoge als lage opgeleiden in de steekproef zijn opgenomen. Hetzelfde geldt voor bestaansonzekerheid, dat een gemiddelde heeft van 2,62. De minimale score is 1 en de maximale score is 5, het gemiddelde ligt dus in het midden. Dit laat een normale verdeling zien. Er kan dus gesteld worden dat de steekproef ongeveer evenveel mensen bevat die een lage mate van bestaansonzekerheid ervaren dan mensen die een hoge mate van bestaansonzekerheid ervaren.

De laatste variabele, de controle variabele leeftijd, is wel redelijk normaal verdeeld. Dit houdt in dat alle leeftijdsgroepen even verdeeld zijn in de steekproef. Er zijn ongeveer evenveel “jongeren” als “ouderen” opgenomen in de steekproef.

BIVARATE CORRELATIES

Tabel 2: Product-momentcorrelaties van de variabelen die zijn opgenomen in de analyse

	1	2	3	4	5	6
1. <i>Gezondheid (goed)</i>	-	.				
2. <i>Opleidingsniveau</i>	.131 **	-				
3. <i>Leeftijd</i>	-.13**	-	-			
4. <i>Niet leefstijl-gerelateerde gezondheidsproblemen</i>	.404**	.092* *	- .242* *	-		
5. <i>Geslacht (man)</i>	-.066*	-.038	- .063* *	-.019		-
6. <i>Bestaansonzekerheid</i>	-.177**	- .180* *	- .096* *	- .074**	.133**	-

*significant bij $p < 0,05$; ** significant bij $p < 0,01$; tweezijdige toets; $N=2465$

Tabel 3: gemiddelden van alle variabelen die zijn opgenomen in de analyse voor gezonde en minder gezonde mensen; $N=2465$

Variabele	Gemiddelde minder gezonde mensen	Gezonde mensen
Opleidingsniveau	3,27	3,78
Bestaansonzekerheid	2,92	2,54
Leeftijd	60,56	54,70
Niet leefstijl-gerelateerde gezondheidsproblemen	74,6%	25,9%
<i>Ja</i>	25,4%	74,1%
<i>Nee</i>		
Geslacht		
<i>Man</i>	63,4%	71%
<i>Vrouw</i>	36,6%	29%

De correlaties tussen de onderzoeksvariabelen zijn te zien in tabel 2. De belangrijkste en opvallendste resultaten uit deze tabel zullen in deze sub-paragraaf worden besproken. Omdat de afhankelijke variabele gezondheid een dummyvariabele is, is in tabel 3 van gezonde en minder gezonde mensen het gemiddelden te vinden van alle variabelen. Deze gegevens zeggen in het geval van een dummyvariabele meer dan correlatie.

De belangrijkste resultaten houden verband met de hypothesen die in het theoriehoofdstuk zijn

opgesteld. Uit tabel 3 blijkt dat gezonde mensen gemiddeld een iets hogere opleiding hebben (3,78) dan minder gezonde mensen (3,27). Het verschil in gemiddeldes is niet heel groot, 0,51, maar er is wel een verschil zichtbaar. Ook ervaren hoger opgeleiden een lager bestaansonzekerheidsgevoel dan lager opgeleiden ($r=-0,180$; $p<0,001$). Bovendien ervaren gezonde mensen gemiddeld minder bestaansonzekerheid (2,92) dan mensen met een mindere gezondheid (2,54).

Verder is het ook opvallend dat bestaansonzekerheid afneemt naarmate de leeftijd stijgt ($r=-0,096$; $p<0,001$). Ook de gezondheid ($r=-0,13$; $p<0,001$) en het opleidingsniveau ($r=-0,252$; $p<0,001$) dalen naarmate de leeftijd stijgt. Het verschil in opleidingsniveau tussen jongeren en ouderen zou er mee te maken kunnen hebben dat er vroeger minder mensen gingen studeren, vooral vrouwen, want die bleven thuis om voor de kinderen te zorgen. Dit betekent dat leeftijd dus een effect lijkt te hebben op het verband tussen opleidingsniveau en gezondheid, die via bestaansonzekerheid zou kunnen lopen. Daarom is het goed dat er in de analyse voor leeftijd wordt gecontroleerd, aangezien die dus een schijnvariabele zou kunnen zijn.

Ook valt het op dat mannen gemiddeld een hoger opleidingsniveau hebben dan vrouwen ($r=-0,038$; $p<0,001$). Opvallend is ook dat mannen zowel vaker gezond zijn (71%) als minder gezond (63,4%), dan vrouwen. Dit zou verklaard kunnen worden doordat er meer mannen in de steekproef zijn opgenomen dan vrouwen.

Als laatste is ook te zien in tabel 2 dat mensen die niet leefstijl-gerelateerde gezondheidsproblemen hebben, een minder goede gezondheid hebben dan mensen die geen niet leefstijl-gerelateerde gezondheidsproblemen hebben ($r=0,404$; $p<0,001$). Van de mensen die minder gezond zijn, heeft 74,1 % niet-leefstijl gerelateerde gezondheidsproblemen. Terwijl 25,9% van de gezonde mensen niet-leefstijl gerelateerde gezondheidsproblemen heeft. Dit is natuurlijk niet erg opvallend omdat het ervaren van een lichamelijke klacht ervoor zorgt dat iemand een minder goede gezondheid ervaart of dit nou een leefstijl-gerelateerde oorzaak heeft of niet. Ook valt op dat mensen met niet leefstijl-gerelateerde gezondheidsproblemen, gemiddeld een minder hoog opleidingsniveau hebben ($r=0,092$; $p<0,001$). Deze twee verbanden wijzen niet op een schijnverband. Ondanks dat het bestaansonzekerheidsgevoel wel groter is bij iemand die niet leefstijl-gerelateerde gezondheidsproblemen heeft ($r=-0,074$; $p<0,001$).

MODELEVALUATIE

Tabel 4: Resultaten van een stapsgewijze logistische regressie-analyse met gezondheid als afhankelijke, opleidingsniveau als onafhankelijke, bestaansonzekerheid als mediatorende variabelen en leeftijd, geslacht en lichamelijke klacht zonder mentale oorzaak als controlevariabele (N=3877)

	Model 1a			Model 2a			Model 3b			Model 4a			VIF
	b (SE)	P*	Odds-ratio	b (SE)	P*	Odds-ratio	b (SE)	P*	Odds-ratio	b (SE)	P*	Odds-ratio	
Constance	.654 (.033)	<.001	1.922	-1.123 (.120)	<.006	0.325	3.709 (0.01)	<0.001	.830 (.489)	.090	2.294		
Opleidingsniveau	.216 (.124)	<.001	1.242	0.161 (0.037)	<0.001	1.175	-0.103 (0.01)	<0.001	.106 (.039)	.006	.753	1.120	
Niet leefstijl-gerelateerde gezondheidsproblemen				2.049 (.120)	<0.001	7.764	-0.171 (0.03)	<0.001	2.015 (.121)	<.001	7.497	1.073	
Leeftijd				-.010 (.004)	<.013	0.990	-0.010 (0.00)	<0.001	-.018 (.004)	<.001	.983	1.162	
Geslacht (man)				-.377 (0.09)	0.002	0.686	0.047 (0.03)	<0.001	-.283 (.122)	.020	0.753	1.022	
Bestaansonzekerheid									-.518 (.070)	<0.001	0.596	1.081	
<i>Hosmer & Lemeshow</i>	2.296	0.682		9.921	0.271				2.811	0.946			
X ² -toets	42.380	<0.001		389.181	<0.001				56.065	<0.001			
<i>Deviance</i>	2390.965			2001.784					1945.718				
<i>Adjusted R Square</i>							0.075						
<i>F</i>							49.986	<0.001					
<i>N</i>	2465			2465			2465		2465				

*Significant bij $p < 0,05$

^a afhankelijke variabele is gezondheid; ^b afhankelijke variabele is bestaansonzekerheid

Modelfit

Er is een logistische regressieanalyse uitgevoerd om de eerder opgestelde hypothesen te kunnen toetsen, zie model 4. Voordat er specifiek gekeken kan worden naar de hypothesen, zal eerst de kwaliteit van de getoetste modellen beoordeeld worden. Hiervoor zal gekeken worden naar de deviance, en de bijbehorende chi-kwadraat, van een model. De deviance-maat geeft aan hoe slecht een model is. Dit houdt dan dus in dat als de deviance afneemt, dit betekent dat het model een betere fit is voor de data en dus minder “slecht” wordt. Ook zal er gekeken worden naar de Hosmer-Lemeshowtoets, deze toets geeft weer hoe goed het model bij de data past. Hiervoor geldt hoe hoger het significantieniveau hoe beter.

Eerste zal het model zonder voorspellers vergeleken worden met model 1. In model 1 wordt gezondheid geschat op basis van opleidingsniveau. De chi-square van model 1 is significant ($X^2(1) = 42.380, p < .001$). Dit houdt in dat model 1 een betere fit heeft dan het lege model en dat model 1 dus beter bij de data past dan het lege model. De Hosmer and Lemeshow Test heeft een score van 2.296 ($df=4, p=0,682$), dit laat zien dat model 1 goed bij de data past.

Het tweede model is het model dat naast opleidingsniveau ook geslacht, leeftijd en niet leefstijl-gerelateerde gezondheidsproblemen als voorspeller voor gezondheid bevat. Model 2 heeft een deviance van 2001.784 ($deviance= 2001.784; X^2(4) = 658,307, p < .001$), dit is een daling ten opzichte van de deviance van model 1 ($deviance=2390.965; X^2(1) = 42.380, p < .001$). Deze daling laat zien dat model 2 dus minder slecht is dan model 1 en dat model 2 dus een betere modelfit is dan model 1. De Hosmer and Lemeshow Test heeft een score van 9.921 ($df=8, p=0.271$), dit laat zien dat het model goed bij de data past.

Naast dat er drie modellen zijn uitgevoerd doormiddel van logistische regressieanalyse, is er ook één model uitgevoerd met lineaire regressie. Dit model schat bestaansonzekerheid op basis van opleidingsniveau en de controlevariabelen geslacht, leeftijd en niet leefstijl-gerelateerde gezondheidsproblemen. Voor dit model wordt er gekeken naar de Rchange en de F-score. De $R^2 adjusted$ van model 3 is 0.075. Inhoudelijk betekent dit, dat 7,5% van de variantie in gezondheid in model 3 wordt verklaard. Deze variantie is significant ($F(4) = 49.986; p < .001$). De verklaarde variantie in gezondheid neemt dus toe, vergeleken met het lege model ($Fchange(4,2460) = 49.986, p < .001$). Deze toename in variantie is significant.

Als laatste zal model 2 vergeleken worden met model 4, het model dat naast opleidingsniveau, geslacht, leeftijd en het hebben van een lichamelijke klacht zonder mentale oorzaak ook bestaansonzekerheid als voorspeller voor gezondheid bevat. Model 4 heeft een deviance van 1945.718 ($deviance=1945.718; X^2(5) = 56.065, p < .001$), dit is een kleine daling ten opzichte van de deviance van model 2. Deze daling laat zien dat model 4 dus een beetje minder slecht is dan model 2 en dat model 4 dus een wat betere modelfit is dan model 2. De Hosmer and Lemeshow Test heeft een score van 2811 ($df=8, p=0.946$), dit laat zien dat het model goed bij de data past.

Assumpties

Bij logistische regressieanalyse zijn er twee assumpties. De eerste is onafhankelijke waarnemingen en de tweede is lineair verband. Lineair verband is moeilijk te controleren en dit is ook geen verplichte eis bij het doen van een logistische regressieanalyse. Daarom zal deze assumptie ook niet verder onderzocht worden. Er wordt dan dus aangenomen dat er aan deze assumptie is voldaan. De andere assumptie, onafhankelijke waarnemingen, houdt in dat de ene waarneming de andere niet mag beïnvloeden. Doordat er in deze steekproef is gecontroleerd voor huishoudhoofden, kan er gesteld worden dat er aan deze assumptie is voldaan. Op deze manier kunnen familieleden en hun zelfde visie elkaar niet beïnvloeden. Ook mag hiervan uit worden gegaan, omdat de steekproef willekeurig is getrokken (*LISS-panel*, z.d.).

Multicollineariteit

Nu er aan beide assumpties voldaan lijkt, moet er nog gekeken worden naar de VIF-scores om iets te kunnen zeggen over de multicollineariteit. De mate van multicollineariteit laat de samenhang tussen de verschillende predictoren zien. Een samenhang tussen de predictoren die te groot is, laat zien dat de predictoren elkaar overlappen in het verklaren van de variantie in gezondheid. De multicollineariteit geeft ook de standaardfouten weer die te groot zijn, waardoor het model onnauwkeuriger wordt. Voor de VIF-waarden geldt, dat er bij een VIF-waarde die groter is dan vier, sprake is van multicollineariteit. Zoals te zien is in tabel 3, is er geen enkele VIF-waarde groter dan vier en is er geen sprake van predictoren die elkaar overlappen of te grote standaardfouten. De mate van multicollineariteit in de analyse is dus goed.

Outliers

Als laatste is er ook nog gekeken naar outliers. Een uitgebreide uitleg hierover is te vinden in bijlage 3. Er zijn verschillende soorten outliers bekeken. Hiervoor is gekeken naar leverage, DfBeta en gestandaardiseerde residuen. Samenvattend zijn er veel outliers gevonden tijdens het bekijken van de verschillende waardes. Er zijn maar 34 cases die op twee manieren een mogelijke outlier zouden kunnen zijn. Na het bekijken van deze afzonderlijke cases is uiteindelijk besloten de cases niet uit de dataset te halen. Hiervoor is gekozen, omdat er geen duidelijke afwijkende waarden waren. Dit zijn waarden die erg afwijken van de waarde ervoor. Ook lijken de waarden geen bijzondere scores te hebben, die de data zouden kunnen beïnvloeden. Als laatste is ook rekening gehouden met de verdeling van de variabele gezondheid, dit zou een redenen kunnen zijn voor de afwijkende waarden van de cases. Om al deze redenen is dus besloten de cases niet te verwijderen.

HYPOTHESE TOETSING

In tabel 4 zijn de resultaten te zien van de logistische regressieanalyse. Aan de hand van deze resultaten kunnen de hypothesen getoetst worden. Hypothese 1 stelt dat hoger opgeleiden een betere gezondheid hebben dan lager opgeleiden. Uit model 2 blijkt dat opleidingsniveau een significant effect heeft op de kans op gezondheid ($b=0,161$; $p<,001$). Dit houdt in dat hoger opgeleiden een grotere kans hebben om gezonder te zijn dan lager opgeleiden. De resultaten steunen dus de hypothese “hoger opgeleiden hebben een betere gezondheid dan lager opgeleiden”.

Hypothese 2 stelt dat hoger opgeleiden gezonder zijn omdat ze minder bestaansonzekerheid ervaren dan laagopgeleiden. Voor deze hypothese moet er ondersteuning gevonden worden voor verschillende verbanden.

Allereerst moeten hoger opgeleiden minder bestaansonzekerheid ervaren dan laagopgeleiden. Model 3 laat zien dat opleidingsniveau een negatief effect heeft op bestaansonzekerheid ($b=-0,103$; $p<,001$). Dit betekent dat iemand met een hoger opleidingsniveau inderdaad een minder grotere kans heeft op het ervaren van bestaansonzekerheid dan iemand met een lager opleidingsniveau. Voor dit verband is dus ondersteuning gevonden.

Ten tweede moeten mensen die meer bestaansonzekerheid ervaren over het algemeen ook een minder goede gezondheid ervaren dan mensen die in mindere mate bestaansonzekerheid ervaren. Model 4 laat zien dat bestaansonzekerheid een significant effect heeft op gezondheid ($b=0,518$; $p<,001$). Dit houdt in dat mensen die veel bestaansonzekerheid ervaren een grotere kans hebben op een goede gezondheid dan mensen die minder bestaansonzekerheid ervaren. Voor dit verband is dus ook ondersteuning gevonden.

Wanneer bestaansonzekerheid wordt toegevoegd aan het bestaande model dat de kans op gezondheid schat is het effect van opleidingsniveau op gezondheid nog wel significant, maar wel iets kleiner geworden vergeleken met model 2 ($b=0,106$; $p<,006$). Dit effect is relatief klein. Wanneer gekeken wordt naar de betrouwbaarheidsintervallen van gezondheid in model 2 (1,092;1,265) en model 4 (1,031;1,199), valt het op dat die elkaar wel overlappen. De betrouwbaarheidsintervallen overlappen elkaar voor 45,5%, dit is veel overlap. Dat houdt dus in dat er niet voldoende evidentie is om te stellen dat de verbanden van elkaar verschillen en dat bestaansonzekerheid dus inderdaad kan verklaren dat laagopgeleiden minder gezond zijn dan hoogopgeleiden. De resultaten steunen hypothese 2 dus niet.

Verder is er ook gecontroleerd voor controle variabelen. Allereerst is uit tabel model 2 op te maken dat het toevoegen van de controle variabelen aan model 1, het effect van opleidingsniveau op gezondheid verkleint. Dit effect gaat van 0,216 naar 0,161. Dit laat zien dat de controlevariabelen dus wel effect hebben op het verband tussen opleidingsniveau en gezondheid. Wel is het effect van opleidingsniveau op gezondheid nog steeds best groot en significant. Ook is uit de tabel op te maken dat ouder worden een negatief effect heeft op gezondheid ($b=-0,01$; $p=-0,001$). Oudere mensen zijn

over het algemeen dus minder gezond dan jongere mensen. Over geslacht kan vastgesteld worden dat mannen een iets grotere kans op een goede gezondheid hebben dan vrouwen ($b=-0,01$; $p=-0,042$). Als laatste is uit de tabel op te maken dat het ervaren van niet leefstijl-gerelateerde gezondheidsproblemen een positief effect heeft op de gezondheid van een persoon ($b=-2,10$; $p=-0,042$). Dit betekent dat wanneer iemand geen niet leefstijl-gerelateerde gezondheidsproblemen ervaart, diegene een betere kans heeft op een goede gezondheid, dan wanneer iemand wel niet leefstijl-gerelateerde gezondheidsproblemen ervaart.

CONCLUSIE & DISCUSSIE

CONCLUSIE

De onderzoeksvraag die centraal stond in dit onderzoek is: “kan bestaansonzekerheid een verklaring bieden voor de lagere gezondheid van laagopgeleiden ten opzichte van hoogopgeleiden?”. Uit dit onderzoek is gebleken dat lager opgeleiden zowel meer bestaansonzekerheid als een slechtere gezondheid ervaren en dat een hogere mate van bestaansonzekerheid leidt tot een minder goede gezondheid. Toch is er geen ondersteuning gevonden voor dat bestaansonzekerheid een verklaring biedt voor het gezondheidsverschil tussen hoog- en laagopgeleiden.

Door het terugkoppelen van de resultaten van dit onderzoek naar het in de inleiding besproken maatschappelijk debat, kan de bijdrage van dit onderzoek aan de opleidingskloof worden vastgesteld. Allereerst wordt in de inleiding gesteld dat het van belang is dat de algemene gezondheid van de samenleving wordt bevorderd. Redenen hiervoor waren een tekort aan zorgmedewerkers, stijgende zorgkosten en een ontworpen samenleving. Uit dit onderzoek is gebleken dat het ervaren van meer bestaansonzekerheid leidt tot een minder goede gezondheid. Het verminderen van bestaansonzekerheid zou daardoor dus de gezondheid van de samenleving kunnen verbeteren. Mensen meer zekerheid bieden, door bijvoorbeeld het minimumloon te verhogen, zou gunstig moeten zijn voor de gezondheid. Dit onderzoek heeft dus bijgedragen aan de vraag hoe de gezondheid verbeterd kan worden door aan te tonen dat bestaansonzekerheid een verklaring biedt voor een minder goede gezondheid.

Ondanks dat bestaansonzekerheid wel een verklaring lijkt te zijn voor een mindere gezondheid, heeft dit onderzoek aangetoond dat het geen verklaring biedt voor het verschil in gezondheid tussen hoog- en laagopgeleiden. Zoals eerder in dit onderzoek al werd benoemd, wordt het verschil in kennis tussen hoog- en laagopgeleiden gezien als de meest aannemelijke verklaring voor het gezondheidsverschil tussen hoog- en laagopgeleiden. De dominante interventie voor het verkleinen van gezondheidsverschillen tussen hoog- en laagopgeleiden is op dit moment dan ook voorlichting geven over een gezonde leefstijl. Dit onderzoek heeft aangetoond dat bestaansonzekerheid geen verklaring biedt voor het verschil tussen hoog- en laagopgeleiden. Deze verklaring is dus uitgesloten. Dit maakt het aannemelijker dat het verschil in kennis, over gezondheid, wel een verklaring biedt voor de gezondheidsverschillen tussen hoog- en laagopgeleiden.

DISCUSSIE

Zoals in de conclusie al werd genoemd is er geen ondersteuning gevonden voor dat bestaansonzekerheid een verklaring biedt voor het gezondheidsverschil tussen hoog- en laagopgeleiden. Wel zijn er een aantal dingen waar rekening mee gehouden moet worden bij het doen van uitspraken over deze resultaten. Zo zijn er in dit onderzoek een aantal methodologische keuzes gemaakt die van invloed kunnen zijn op de resultaten.

De eerste methodologische keuze waar rekening mee gehouden moet worden is de manier waarop de afhankelijke variabele gezondheid is geoperationaliseerd. Doordat er uiteindelijk een logistische regressieanalyse is uitgevoerd in dit onderzoek moest de originele variabele gezondheid omgezet worden in een dummyvariabele. Door deze dummyvariabele was de verdeling tussen gezonde en minder gezonde mensen erg scheef, er zijn veel meer gezonde mensen dan minder gezonde mensen opgenomen in de steekproef. Dit houdt in dat de resultaten ook vertekend kunnen zijn doordat er veel meer gezonde mensen zijn meegenomen in de analyse. Als de verdeling eerlijker zou zijn geweest, zou met meer zekerheid iets gezegd kunnen worden over de resultaten, omdat nu één groep over verantwoord is.

In dit onderzoek wordt onderzocht of opleiding een effect heeft op bestaansonzekerheid, dus of wanneer iemand een lage opleiding heeft diegene ook meer bestaansonzekerheid zal ervaren. Dit impliceert dat wanneer meer mensen een hoge opleiding doen er minder bestaansonzekerheid zal zijn. Dit is ook zo, maar hierbij wordt niet meegenomen dat het verschil in opleidingsniveau ook weer kan ontstaan door bestaansonzekerheid. Doordat iemands ouders veel bestaansonzekerheid ervaren is het aannemelijk dat ze ook niet genoeg geld hebben om hun kind te laten studeren. Ook zal het kind minder gestimuleerd worden te studeren als de ouders dit zelf ook niet hebben gedaan. Dit houdt in dat het zo zou kunnen zijn dat wanneer er meer bestaanszekerheid wordt gecreëerd, er ook meer mensen hoogopgeleid zullen zijn. Dit onderzoek kijkt niet naar het effect van bestaanszekerheid op opleidingsniveau. Daarom kan er niks gezegd worden over de wisselwerking tussen opleidingsniveau en bestaansonzekerheid.

Ook generaliseerbaarheid is een mogelijke beperking van dit onderzoek. Het onderzoek is gedaan doormiddel van al bestaande data van het LISS-panel, dit panel bevat alleen Nederlandse huishoudens. Daardoor bevat dit onderzoek ook alleen maar Nederlandse data. Dit maakt het onderzoek alleen representatief voor Nederland en niet voor andere landen. Nederland is bijvoorbeeld een heel ander land dan de Verenigde Staten (VS). Zo is de zorg in Nederland veel betaalbaarder dan in de VS (Wilson et al., 2020). Dit heeft te maken met dat Nederland een verzorgingsstaat is en de VS niet. In de VS moet medische zorg bijvoorbeeld grotendeels zelf betaald worden. Amerikanen kunnen zich wel laten verzekeren, maar dat is erg duur (tussen de 500 en 800 euro per maand). Daarom kiezen veel Amerikanen ervoor om zich niet te laten verzekeren. Ook voor hoogopgeleiden, die meer verdienen, kunnen dure zorgkosten dan dus voor bestaansonzekerheid zorgen. Een grote medische

ingreep kost in de VS heel veel geld, terwijl dat in Nederland onder de verzekering valt. Inhoudelijk wil dit zeggen dat bestaansonzekerheid bij zowel hoog- als laagopgeleiden een rol speelt in de VS en daarom ook geen verklaring zou kunnen bieden voor de gezondheidsverschillen tussen hoog- en laagopgeleiden

Na de gevonden resultaten van dit onderzoek zou het interessant kunnen zijn om vervolgonderzoek te doen. Dit onderzoek heeft aangetoond dat bestaansonzekerheid geen verklaring biedt voor de gezondheidsverschillen tussen hoog- en laagopgeleiden. Dit wil zeggen dat de veel gebruikte verklaring, het verschil in kennis, inderdaad een verklaring zou kunnen bieden voor het verschil in gezondheid tussen hoog- en laagopgeleiden. Op dit moment is de dominante interventie voor het verkleinen van gezondheidsverschillen tussen hoog- en laagopgeleiden, voorlichting geven over een gezonde leefstijl. Nu zou het interessant zijn om te kijken hoeveel van het verschil tussen hoog- en laagopgeleiden toe te schrijven is aan hetgeen wat ze op school leren. Zo kan gekeken worden waar het verschil in kennis precies vandaan komt.

Ook is er een beleidsimplicatie die gedaan kan worden naar aanleiding van de bevindingen uit dit onderzoek. Zoals in de inleiding al werd vermeld is er een tekort aan zorgmedewerkers, stijgen de zorgkosten en ontwricht te veel ziekte de samenleving. Om deze problemen, deels, te verhelpen moet er gekeken worden naar hoe de gezondheid bevorderd kan worden. De bevindingen van dit onderzoek hebben uitgewezen dat bestaansonzekerheid geen verklaring biedt voor de gezondheidsverschillen tussen laag- en hoogopgeleiden. Wel ervaren mensen die meer bestaansonzekerheid ervaren ook een minder goede gezondheid, dan mensen die in mindere mate bestaansonzekerheid ervaren.

Bestaansonzekerheid zou dus wel kunnen helpen in het bevorderen van de gezondheid. Hiervoor zou er gekeken kunnen worden naar manieren om de bestaansonzekerheid te verkleinen

LITERATUURLIJST

- Bussi, M., & Pareliussen J. (2017). Back to Basics- Literacy Proficiency, Immigration and Labour Market Outcomes in Sweden. *Social Policy & Administration*, 51(4), 676-696. <https://doi-org.proxy-ub.rug.nl/10.1111/spol.12319>
- Bruggink, J. W. (2009). Ontwikkelingen in (gezonde) levensverwachting naar opleidingsniveau. *Bevolkingstrends*, 57(4), 71-74.
- De Groot, N. (2022, 14 november). Zorgpremies 2023 bekend: forse stijging meest gekozen polissen. *AD.nl*. <https://www.ad.nl/privacy-gate/accept-tcf2?redirectUri=%2Fgeld%2Fzorgpremies-2023-bekend-forse-stijging-meest-gekozen-polissen%7Ea6eae0eb%2F&authId=e6f5bb13-3206-4c73-9100-e0683b09f249&referrer=https%3A%2F%2Fwww.google.com%2F&referer=https%3A%2F%2Fwww.google.com%2F>
- Gökbulut, B., Yeniasur, M. (2018). Analysis of children's interest in books and their reading levels depending on the education status of family. *Quality & Quantity*, 52(1), 235-245. <https://doi-org.proxy-ub.rug.nl/10.1007/s11135-017-0608-2>
- Haas, S. A., & Fosse, N. E. (2008). Health and the educational attainment of adolescents: evidence from the NLSY97. *Journal of health and social behavior*, 49(2), 178–192. <https://doi.org/10.1177/002214650804900205>
- Hoeymans, N. (2013). Naar een nieuwe definitie van gezondheid. *Tsg- Tijdschrift Voor Gezondheidswetenschappen*. <https://doi.org/10.1007/s12508-013-0044-1>
- Hosseini, H. (2015). Aging and the Rising Costs of Healthcare in the United States: Can There be a Solution? *Ageing International*, 40(3), 229-247. <https://doi-org.proxy-ub.rug.nl/10.1007/s12126-014-9209-8>
- Jian, J. (2023). Relationship between social cohesion and basic public health services utilization among Chinese internal migrants: a perspective of socioeconomic status differentiation. *Health Sociology Review*, 32(2), 179-197. <https://doi-org.proxy-ub.rug.nl/10.1080/14461242.2022.2112410>
- Kasl S. V. (1984). Stress and health. *Annual review of public health*, 5, 319–341. <https://doi.org/10.1146/annurev.pu.05.050184.001535>
- Kerssemakers, R., Van Meerten, R., Noorlander, E., & Vervaeke, H. K. (2009). *Drugs en alcohol; Gebruik, misbruik en verslaving*. <https://doi.org/10.1007/978-90-313-6554-8>
- Kronenfeld J.J. (1999). Gender and Health Status. In *Handbook of the Sociology of Gender* (pp.459-481). <https://doi.org/10.1007/0-387-326218-5>
- Lagaay, A. M. (1999). Gezondheid en ziekte op oudere leeftijd. *Geriatrische tandheelkunde: De problematiek van ouder worden en mondgezondheid*, 33-43.

- Launoy G. (2023). Social inequalities in health. For an evidence-based public policy. *Journal of visceral surgery*, S1878-7886(23)00075-9. Advance online publication. <https://doi.org/10.1016/j.jviscsurg.2023.04.009>
- Leigh, J.P. (1983). Direct and Indirect Effects of Education on Health. *Social Science & Medicine*, 17(4), 227-234. [https://doi-org.proxy-ub-rug.nl/10.1016/0277-9526\(83\)90120-X](https://doi-org.proxy-ub-rug.nl/10.1016/0277-9526(83)90120-X)
- Link, B. G., & Phelan, J. C. (2002). McKeown and the idea that social conditions are fundamental causes of disease. *American journal of public health*, 92(5), 730–732. <https://doi.org/10.2105/ajph.92.5.730>
- LISS Panel (2022). Centerdata, Tilburg University. <https://www.lissdata.nl/>
- McGill, N. (2016). Social Determinants of Health: Education Attainment Linde to Heath Throughout Life Span. *American Journal of Public Health*, 106(10)m 1719. <https://doi-org.proxy-ub-rug.nl/102105/AJPH.2016.303408>
- Ministerie van Algemene Zaken. (2022, 29 december). *Dalende koopkracht: oorzaken, gevolgen en maatregelen*. Koopkracht | Rijksoverheid.nl. <https://www.rijksoverheid.nl/onderwerpen/koopkracht/dalende-koopkracht-oorzaken-gevolgen-en-maatregelen>
- NOS. (2022b, januari 20). Onderzoek: tekort aan zorgpersoneel op lange termijn alleen maar groter. *NOS*. <https://nos.nl/artikel/2413851-onderzoek-tekort-aan-zorgpersoneel-op-lange-termijn-alleen-maar-groter>
- O'Donnell, S. (2020). 'Your wealth is your health': the fundamental causes of inequalities in diabetes management outcomes: a qualitative analysis. *Sociology of health & illness*, 42(7), 1626–1641. <https://doi.org/10.1111/1467-9566.13156>
- Pflanz, M & Keupp, H. (1977). A sociological perspective on concepts of disease. *International Social Science Journal*, 29(3), 386
- Rademaker, A.L., & Holland, L. R. (2021). Subjective experience of health, personal health resources and beliefs: supporting a lifeworld approach to social work in youth health. *European Journal of Social Work*, 24(2), 344-357. <https://doi-org.proxy-ub-rug.nl/10.1080/13691457.2020.1760794>
- Rebechi, A., & Rohde, N. (2022). Economic insecurity, racial anxiety, and right-wing populism. *Review of Income and Wealth*. <https://doi.org/10.1111/roiw.12599>
- Ring Jr., R J. (1985). Variability of Inflation and Income across Income Classes. *Social Science Quarterly (University of Texas Press)*, 66(1), 203-209
- Standing, G. (2011). The Precariat: the new dangerous class. London: Bloomsbury, 198 pp. *Profesorado: Revista de curriculum y formación del profesorado*, 16(1), 416–417. <https://dialnet.unirioja.es/servlet/articulo?codigo=7235734>
- Stronks, K., Crielaard, L., Galenkamp, H., & Nicolaou, M. (2019). Ongelijke kansen op gezondheid: Hoe krijgen we meer grip op dit complexe verschijnsel? In W. van de Werfhorst & E. van

- Hest (red.), *Gelijke kansen in de stad* (pp. 142-160). Amsterdam: Amsterdam University Press.
- Whitehead, M., & Dahlgren, G. (1991). What can be done about inequalities in health?. *Lancet (London, England)*, 338(8774), 1059–1063. [https://doi.org/10.1016/0140-6736\(91\)91911-d](https://doi.org/10.1016/0140-6736(91)91911-d)
- Wilson, M., Guts A., Waddell, K., Lavis, J., Reid, R., & Evans, C (2020). The impacts of accountable care organizations on patient experience, health outcomes and costs: a rapid review. *Journal of Health services Research & Policy*, 25(2), 130-138. <https://doi-org.proxy-ub.rug.nl/10.1177/1355819620913141>
- You-Wei Yan, Rong-Mao Lin, Yan-Kui Su & Ming Yan Liu (2018). The Relationship between Adolescent Academic Stress and Sleep Quality: A Multiple Mediation Model. *Social Behavior & Personality: An International Journal* 46(1), 63-78. <https://doi-org.proxy-ub.rug.nl/10.2224/sbp.6530>
- Zhou, Z., Fang, Y., Zhou, Z., Li, D., Wang, D., Li, Y., Lu, L., Gao, J., & Chen, G (2017). Assessing Income-Related Health Inequality and Horizontal Inequity in China. *Social Indicators Research*, 132(1), 241-256. <https://doi-org.proxy-ub.rug.nl/10.1007/s11205-015-1221-1>

BIJLAGE 1

SELECTIES UITGEVOERD OP DATA

In dit gedeelte van de bijlage zijn alle selecties te vinden die zijn uitgevoerd op het databestand. Zo is er geselecteerd op huishoudhoofden en zijn respondenten die niet op alle voor dit onderzoek belangrijke vragen hebben geantwoord, uit het databestand gefilterd.

Syntax van de selectie op huishoudhoofden:

```
DATASET ACTIVATE DataSet1.  
USE ALL.  
COMPUTE filter_$=(positie=1).  
VARIABLE LABELS filter_$ 'positie=1 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_$ (f1.0).  
FILTER BY filter_$.  
EXECUTE.
```

Syntax van de selectie van het filteren van de missing waarden:

```
DATASET ACTIVATE DataSet1.  
REGRESSION  
/MISSING LISTWISE  
/STATISTICS COEFF OUTS R ANOVA  
/CRITERIA=PIN(.05) POUT(.10)  
/NOORIGIN  
/DEPENDENT ch13g004  
/METHOD=ENTER geslacht ch13g018 leeftijd  
/METHOD=ENTER oplcat  
/METHOD=ENTER Bestaansonzekerheid  
/SAVE RESID.
```

```
RECODE RES_1 (SYSMIS=0) (ELSE=1) INTO RES_1_Recode.  
EXECUTE.
```

```
USE ALL.  
COMPUTE filter_$=(RES_1_Recode=1).  
VARIABLE LABELS filter_$ 'RES_1_Recode=1 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_$ (f1.0).  
FILTER BY filter_$.  
EXECUTE.
```

CONTRALE SAMENHANG VAN DE SCHAALVARIABLE BESTAANSONZEKERHEID

Dit gedeelte van de bijlage laat zien dat de cronbach's alpha van de schaalvariabele bestaansonzekerheid 0,773 is. Dit is goed, een score hoger dan 0,70 wijst namelijk op een grote mate van interne consistentie en op een hoge mate van betrouwbaarheid. Dit houdt in dat er dus een schaalvariabele van deze losse vragen gemaakt kan worden.

Syntax:

```
RELIABILITY
/VARIABLES=ci13f252_new Ci13f258_new Ci13f244_NEW Ci13f356_new
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/SUMMARY=TOTAL.
```

Reliability Statistics

Cronbach's Alpha	N of Items
.773	4

Case Processing Summary

Cases		N	%
		Valid	2342
	Excluded ^a	123	5.0
	Total	2465	100.0

a. Listwise deletion based on all variables in the procedure.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
ci13f252_new	7.6917	6.361	.646	.684
Ci13f258_new	7.4821	7.786	.516	.757
Ci13f244_NEW	7.8646	6.478	.686	.670
Ci13f356_new	7.9761	5.078	.561	.768

FREQUENTIEVERDELING, DESCRIPTIVES EN OPERATIONALISATIES VAN DE VARIABELEN

In dit gedeelte van de bijlage zijn frequentieverdelingen en descriptives te vinden van alle variabelen. Zowel van de oorspronkelijke variabelen als van de uiteindelijke variabelen. Ook de operationalisaties die zijn uitgevoerd op de variabelen zijn in dit gedeelte weer te vinden. Dit is gedaan voor alle variabelen die in de analyse voortkomen, dus alle variabelen waar de hiervoor uitgelegde selecties op toegepast zijn.

AFHANKELIJKE VARIABELE GEZONDHEID**De oorspronkelijke items die de afhankelijke variabele gezondheid uiteindelijk meten**

De verdeling lijkt redelijk normaal verdeeld. Het gemiddelde (3.04) ligt redelijk in het midden van de vijf antwoordmogelijkheden, dus van het minimum en het maximum. Ook aan het histogram is te zien dat de verdeling normaal verdeeld is, alleen komt categorie 3 “goed” wel veel meer voor dan de andere categorieën. De verdeling is verder wel mooi normaal verdeeld.

Syntax:

```
DATASET ACTIVATE DataSet1.
DESCRIPTIVES VARIABLES=ch13g004
/STATISTICS=MEAN STDDEV MIN MAX.
```

```
FREQUENCIES VARIABLES=ch13g004
/ORDER=ANALYSIS.
```

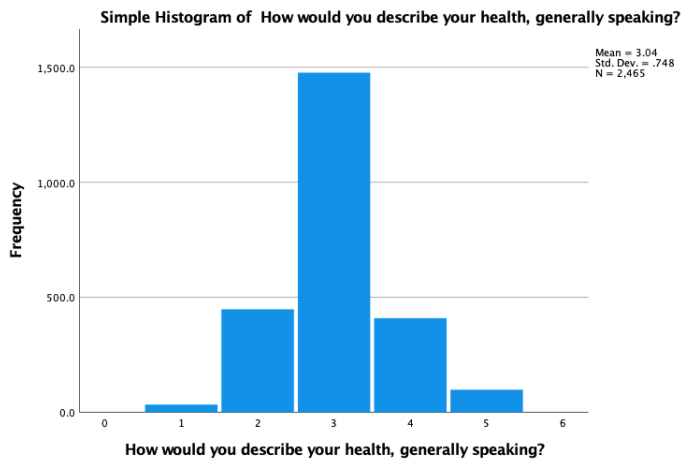
* Chart Builder.

```
GGRAPH
/GRAPHDATASET NAME="graphdataset" VARIABLES=ch13g004 MISSING=LISTWISE
REPORTMISSING=NO
/GRAPHSPEC SOURCE=INLINE.
BEGIN GPL
SOURCE: s=userSource(id("graphdataset"))
DATA: ch13g004=col(source(s), name("ch13g004"))
GUIDE: axis(dim(1), label("How would you describe your health, generally speaking?"))
GUIDE: axis(dim(2), label("Frequency"))
GUIDE: text.title(label("Simple Histogram of How would you describe your health, generally ",
"speaking?"))
ELEMENT: interval(position(summary.count(bin.rect(ch13g004))), shape.interior(shape.square))
END GPL.
```

		Descriptive Statistics				
		N	Minimum	Maximum	Mean	Std. Deviation
ch13g004	How would you describe your health, generally speaking?	2465	1	5	3.04	.748
Valid N (listwise)		2465				

ch13g004 How would you describe your health, generally speaking?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 poor	33	1.3	1.3	1.3
	2 moderate	448	18.2	18.2	19.5
	3 good	1477	59.9	59.9	79.4
	4 very good	409	16.6	16.6	96.0
	5 excellent	98	4.0	4.0	100.0
Total		2465	100.0	100.0	



De uiteindelijk gebruikte dummyvariabele van gezondheid

Van de variabele gezondheid is een dummyvariabele gemaakt omdat er uiteindelijk is gekozen voor een logistische regressie. De reden hiervoor is te vinden in de analyse-opzet en in bijlage 2 en 3, waar verder wordt ingegaan op de geschonden assumpties van lineaire regressie. De dummyvariabele is scheef verdeeld. Er zijn veel meer mensen met een goede gezondheid dan dat er mensen zijn met een minder goede gezondheid. Dit heeft te maken met dat de oorspronkelijke vraag 5 antwoordcategorieën had.

Syntax:

```
RECODE ch13g004 (1 thru 2=0) (3 thru 5=1) INTO Gezondheid_dummy.  
EXECUTE.
```

```
DESCRIPTIVES VARIABLES=Gezondheid_dummy  
/STATISTICS=MEAN STDDEV MIN MAX.
```

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

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Gezondheid_dummy	2465	.00	1.00	.8049	.39638
Valid N (listwise)	2465				

Statistics		
Gezondheid_dummy		
N	Valid	2465
	Missing	0

		Gezondheid_dummy			
		Frequenc y	Percent	Valid Percent	Cumulative Percent
Valid	.00	481	19.5	19.5	19.5
	1.00	1984	80.5	80.5	100.0
Total		2465	100.0	100.0	

ONAFHANKELIJKE VARIABELE OPLEIDINGSNIVEAU

De variabele is niet helemaal normaal verdeeld. De verdeling is wat linksscheef en de categorie havo/vwo wat ondervertegenwoordigd, verder ligt het gemiddelde wel redelijk in het midden van het minimum en het maximum, dus dit lijkt verder geen probleem te zijn.

Syntax:

```
DESCRIPTIVES VARIABLES=oplcat
  /STATISTICS=MEAN STDDEV MIN MAX.
```

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

* Chart Builder.

```
GGRAPH
  /GRAPHDATASET NAME="graphdataset" VARIABLES=oplcat COUNT()[name="COUNT"]
  MISSING=LISTWISE
  REPORTMISSING=NO
  /GRAPHSPEC SOURCE=INLINE.
```

```
BEGIN GPL
  SOURCE: s=userSource(id("graphdataset"))
  DATA: oplcat=col(source(s), name("oplcat"), unit.category())
  DATA: COUNT=col(source(s), name("COUNT"))
  GUIDE: axis(dim(1), label("Level of education in CBS (Statistics Netherlands) categories"))
  GUIDE: axis(dim(2), label("Count"))
  GUIDE: text.title(label("Simple Histogram Count of Level of education in CBS (Statistics ",
    "Netherlands) categories"))
  SCALE: linear(dim(2), include(0))
  ELEMENT: interval(position(oplcat*COUNT), shape.interior(shape.square))
END GPL.
```

	Descriptive Statistics				
	N	Minimum	Maximum	Mean	Std. Deviation
oplcat Level of education in CBS (Statistics Netherlands) categories	2465	1	6	3.68	1.528
Valid N (listwise)	2465				

Statistics

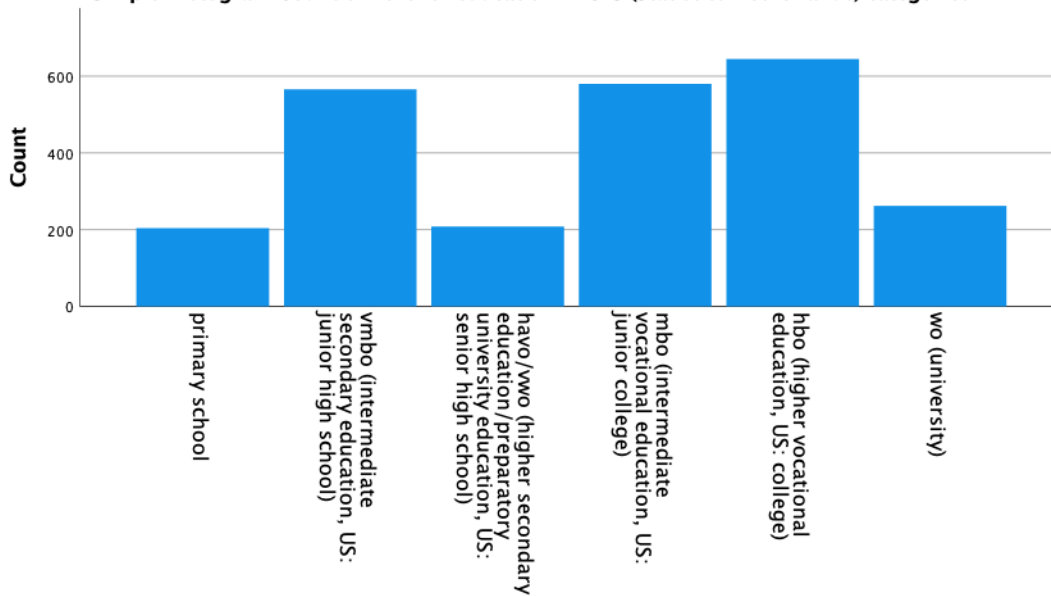
oplcats Level of education in CBS (Statistics Netherlands) categories

N	Valid	2465
	Missing	0

oplcats Level of education in CBS (Statistics Netherlands) categories

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 primary school	204	8.3	8.3	8.3
	2 vmbo (intermediate secondary education, US: junior high school)	566	23.0	23.0	31.2
	3 havo/vwo (higher secondary education/preparatory university education, US: senior high school)	208	8.4	8.4	39.7
	4 mbo (intermediate vocational education, US: junior college)	580	23.5	23.5	63.2
	5 hbo (higher vocational education, US: college)	645	26.2	26.2	89.4
	6 wo (university)	262	10.6	10.6	100.0
	Total	2465	100.0	100.0	

Simple Histogram Count of Level of education in CBS (Statistics Netherlands) categories



Level of education in CBS (Statistics Netherlands) categories

MEDIATIE-VARIABLE BESTAANSONZEKERHEID

Oorspronkelijke items

Syntax:

DESCRIPTIVES VARIABLES=ci13f244 ci13f252 ci13f258 ci13f356
/STATISTICS=MEAN STDDEV MIN MAX.

FREQUENCIES VARIABLES=ci13f252 ci13f244 ci13f258 ci13f356
/ORDER=ANALYSIS.

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
ci13f244 Can you indicate, on a scale from 0 to 10, how hard or how easy it is for you to live off your income?	2465	0	10	6.46	2.002
ci13f252 How would you describe the financial situation of your household at this moment?	2465	1	5	3.33	1.008
ci13f258 Think about the coming 12 months. Do you think that the expenditure of your household will be:	2375	1	5	3.13	.791
ci13f356 How much of a burden is it for you to pay for the living costs of you household	2464	1	7	3.19	1.742
Valid N (listwise)	2374				

ci13f252 How would you describe the financial situation of your household at this moment?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 we are accumulating debts	98	4.0	4.0	4.0
	2 we are somewhat eating into savings	468	19.0	19.0	23.0
	3 we are just managing to make ends meet	650	26.4	26.4	49.3
	4 we have a little bit of money to spare	1030	41.8	41.8	91.1
	5 we have a lot of money to spare	219	8.9	8.9	100.0
	Total	2465	100.0	100.0	

ci13f244 Can you indicate, on a scale from 0 to 10, how hard or how easy it is for you to live off your income?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 very hard	32	1.3	1.3	1.3
	1 1	20	.8	.8	2.1
	2 2	64	2.6	2.6	4.7
	3 3	84	3.4	3.4	8.1
	4 4	144	5.8	5.8	14.0
	5 5	340	13.8	13.8	27.7
	6 6	417	16.9	16.9	44.7
	7 7	564	22.9	22.9	67.5
	8 8	520	21.1	21.1	88.6
	9 9	152	6.2	6.2	94.8
	10 very easy	128	5.2	5.2	100.0
	Total	2465	100.0	100.0	

ci13f258 Think about the coming 12 months. Do you think that the expenditure of your household will be:

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 much higher than the income	66	2.7	2.8	2.8
	2 higher than the income	342	13.9	14.4	17.2
	3 approximately equal to the income	1262	51.2	53.1	70.3
	4 lower than the income	634	25.7	26.7	97.0
	5 much lower than the income	71	2.9	3.0	100.0
	Total	2375	96.3	100.0	
Missing	99 I don't know	90	3.7		
	Total	2465	100.0		

ci13f356 How much of a burden is it for you to pay for the living costs of you household

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 1 not at all a burden	591	24.0	24.0	24.0
	2 2	425	17.2	17.2	41.2
	3 3	346	14.0	14.0	55.3
	4 4	438	17.8	17.8	73.1
	5 5	414	16.8	16.8	89.9
	6 6	189	7.7	7.7	97.5
	7 7 a heavy burden	61	2.5	2.5	100.0
	Total	2464	100.0	100.0	
Missing	System	1	.0		
	Total	2465	100.0		

Operationalisaties:

Om een schaalvariabele te kunnen maken zijn de volgende items omgedraaid: “kunt u op een schaal van 0 tot 10 aangeven hoe moeilijk of makkelijk u kunt rondkomen van uw inkomen?”, “hoe zou u op dit moment de financiële situatie van uw huishouden beschrijven?” en “kijk naar de komende 12 maanden. Denkt u dan dat de uitgaven van uw huishouden zijn: (hierop volgen 5 keuzemogelijkheden)”. Ook zijn alle variabelen zo bewerkt dat ze allemaal uit 5 categorieën bestaan.

Syntax operationlisaties;

DATASET ACTIVATE DataSet1.

RECODE ci13f252 (1=5) (2=4) (3=3) (4=2) (5=1) INTO ci13f252_new.

EXECUTE.

RECODE ci13f258 (1=5) (2=4) (3=3) (4=2) (5=1) (99=SYSMIS) INTO Ci13f258_new.

EXECUTE.

RECODE ci13f244 (1 thru 2=5) (3 thru 4=4) (5 thru 6=3) (7 thru 8=2) (9 thru 10=1) INTO

Ci13f244_NEW.

EXECUTE.

RECODE ci13f356 (3=2) (4=3) (5=4) (1 thru 2=1) (6 thru 7=5) INTO Ci13f356_new.

EXECUTE.

Syntax verdelingen:

*FREQUENCIES VARIABLES=ci13f252_new Ci13f258_new Ci13f244_NEW Ci13f356_new

Bestaansonzekerheid

/STATISTICS=STDDEV MINIMUM MEAN MEDIAN

/ORDER=ANALYSIS.

		Statistics				
		ci13f252_ne	Ci13f258_ne	Ci13f244_NE	Ci13f356_ne	Bestaansonzekerheid
		w	w	W	w	
N	Valid	2465	2375	2433	2464	2465
	Missing	0	90	32	1	0
Mean		2.6738	2.8728	2.4871	2.4058	2.6153
Median		2.0000	3.0000	2.0000	2.0000	2.5000
Std. Deviation		1.00804	.79106	.93538	1.41825	.83481
Minimum		1.00	1.00	1.00	1.00	1.00

		ci13f252_new			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	219	8.9	8.9	8.9
	2.00	1030	41.8	41.8	50.7
	3.00	650	26.4	26.4	77.0
	4.00	468	19.0	19.0	96.0
	5.00	98	4.0	4.0	100.0
Total		2465	100.0	100.0	

Ci13f356_new

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	1016	41.2	41.2	41.2
	2.00	346	14.0	14.0	55.3
	3.00	438	17.8	17.8	73.1
	4.00	414	16.8	16.8	89.9
	5.00	250	10.1	10.1	100.0
	Total	2464	100.0	100.0	
Missing	System	1	.0		
Total		2465	100.0		

Bestaansonzekerheid

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	41	1.7	1.7	1.7
	1.25	61	2.5	2.5	4.1
	1.33	3	.1	.1	4.3
	1.50	128	5.2	5.2	9.5
	1.67	7	.3	.3	9.7
	1.75	284	11.5	11.5	21.3
	2.00	302	12.3	12.3	33.5
	2.25	237	9.6	9.6	43.1
	2.33	6	.2	.2	43.4
	2.50	245	9.9	9.9	53.3
	2.67	11	.4	.4	53.8
	2.75	237	9.6	9.6	63.4
	3.00	239	9.7	9.7	73.1
	3.25	175	7.1	7.1	80.2
	3.33	16	.6	.6	80.8
	3.50	155	6.3	6.3	87.1
	3.67	9	.4	.4	87.5
	3.75	104	4.2	4.2	91.7
	4.00	90	3.7	3.7	95.3
	4.25	43	1.7	1.7	97.1
	4.33	14	.6	.6	97.6
	4.50	24	1.0	1.0	98.6
	4.67	7	.3	.3	98.9
	4.75	15	.6	.6	99.5
	5.00	12	.5	.5	100.0
	Total		2465	100.0	100.0

Schaalvariabele bestaansonzekerheid

Van de hierboven getoonde items is een schaalvariabele gemaakt. Al deze items zijn hiervoor samengevoegd. De verdeling van deze variabele is niet helemaal normaal verdeeld. De verdeling is iets rechtsscheef. Er zijn meer mensen die echt laag scores op bestaansonzekerheid dan dat er mensen zijn die echt hoog scores op bestaansonzekerheid. Wel ligt het gemiddelde redelijk in het midden van het minimum en maximum. Dit houdt in dat de meeste mensen ergens rond het gemiddelde scoren, dit is ook te zien aan het histogram. Samenvattend lijkt de verdeling wel redelijk normaal en zijn er ongeveer even veel mensen die onder het gemiddelde scoren als dat er mensen zijn die boven het gemiddelde scoren.

Syntax:

```
DATASET ACTIVATE DataSet1.  
COMPUTE  
Bestaansonzekerheid=MEAN.3(CI13F252_NEW,CI13F258_NEW,CI13F244_NEW,CI13F356_NEW).  
EXECUTE.
```

```
DESCRIPTIVES VARIABLES=Bestaansonzekerheid  
/STATISTICS=MEAN STDDEV MIN MAX.
```

```
DATASET ACTIVATE DataSet2.  
DATASET ACTIVATE DataSet2.  
FREQUENCIES VARIABLES=Bestaansonzekerheid  
/ORDER=ANALYSIS.
```

* Chart Builder.

```
GGRAPH  
/GRAPHDATASET NAME="graphdataset" VARIABLES=Bestaansonzekerheid  
MISSING=LISTWISE REPORTMISSING=NO  
/GRAPHSPEC SOURCE=INLINE.  
BEGIN GPL  
SOURCE: s=userSource(id("graphdataset"))  
DATA: Bestaansonzekerheid=col(source(s), name("Bestaansonzekerheid"))  
GUIDE: axis(dim(1), label("Bestaansonzekerheid"))  
GUIDE: axis(dim(2), label("Frequency"))  
GUIDE: text.title(label("Simple Histogram of Bestaansonzekerheid"))  
ELEMENT: interval(position(summary.count(bin.rect(Bestaansonzekerheid))),  
shape.interior(shape.square))  
END GPL.
```

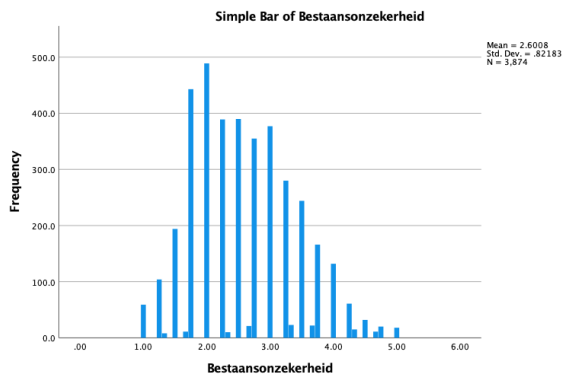
Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Bestaansonzekerheid	2465	1.00	5.00	2.6153	.83481
Valid N (listwise)	2465				

Statistics

Bestaansonzekerheid

N	Valid	2465
	Missin g	0



CONTROLE VARIABLE GEVLACHT

Hieruit is op te merken dat de standaarddeviatie, die 0,5 is, kleiner is dan het gemiddelde van 1,51. Dit zegt niet heel veel omdat er maar 2 antwoordmogelijkheden zijn. 69,5% van de respondenten heeft man geantwoord en 30,5% vrouw. Dit houdt in dat er meer mannen voorkomen in de spreekproef dan vrouwen.

FREQUENCIES VARIABLES=geslacht
/ORDER=ANALYSIS.

DESCRIPTIVES VARIABLES=geslacht
/STATISTICS=MEAN STDDEV MIN MAX.

Statistics

geslacht Gender

N	Valid	2465
	Missing	0

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	1714	69.5	69.5	69.5
	2 Female	751	30.5	30.5	100.0
Total		2465	100.0	100.0	

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
geslacht Gender	2465	1	2	1.30	.460
Valid N (listwise)	2465				

CONTROLE VARIABEL GESLACT

De controle variabele leeftijd is redelijk normaal verdeeld, iets linksscheef. Er komen bijna evenveel jongeren voor in de steekproef dan dat er ouderen voorkomen. Er komen vooral veel 50-70-jarigen voor in de steekproef. Wel komen er veel minder 80+ mensen voor in de steekproef. Dit maakt dat er ongeveer evenveel ouderen in de steekproef voorkomen dan jongeren. De gemiddelde leeftijd ligt ook redelijk in het midden van het minimum en het maximum.

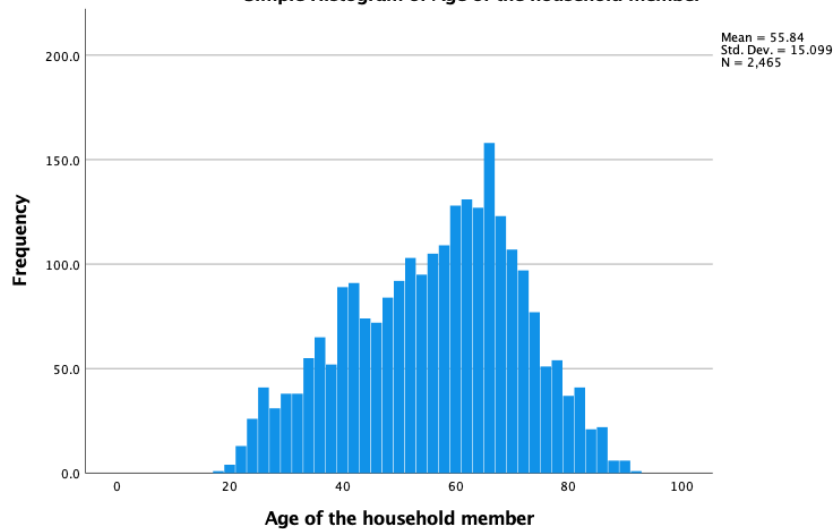
```
FREQUENCIES VARIABLES=leeftijd
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
/ORDER=ANALYSIS.
```

Statistics

leeftijd Age of the household member

N	Valid	2465
	Missing	0
Mean		55.84
Std. Deviation		15.099
Minimum		18
Maximum		92

Simple Histogram of Age of the household member



CONTROLEVARIABLE CHRONISCHE AANDOENING

64,4% van de respondenten heeft nee geantwoord en 35,4% ja. De verdeling is dus wat scheef. Opzich is dit niet heel gek omdat er maar 2 antwoordmogelijkheden zijn, dit hoeft daarom geen probleem te zijn, want er hebben nog steeds wel 872 respondenten ja geantwoord.

```
DESCRIPTIVES VARIABLES=ch13g018
/STATISTICS=MEAN STDDEV MIN MAX.
FREQUENCIES VARIABLES=ch13g018
/ORDER=ANALYSIS.
```

Statistics

leeftijd Age of the household member

N	Valid	2465
	Missing	0
Mean		55.84
Std. Deviation		15.099
Minimum		18
Maximum		92

ch13g018 Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 yes	872	35.4	35.4	35.4
	2 no	1593	64.6	64.6	100.0
	Total	2465	100.0	100.0	

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ch13g018 Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident?	2465	1	2	1.65	.478
Valid N (listwise)	2465				

BIJLAGE 2

BIVARATE ANALYSE.

Een bivariate analyse is uitgevoerd om de onderlinge verbanden tussen twee variabelen te bepalen. Hierdoor kunnen er uitspraken gedaan worden over die onderlinge verbanden. Zie volgende bladzijde voor de output. De belangrijkste resultaten worden besproken in sub paragraaf 4.1.2.

Syntax:

CORRELATIONS

/VARIABLES=geslacht leeftijd oplcat Gezondheid_dummy Bestaansonzekerheid ch13g018

/PRINT=TWOTAIL NOSIG FULL

/MISSING=PAIRWISE

Correlations

		geslacht Gender	leeftijd Age of the househol d member	oplc at Level of education in CBS (Statistics Netherlan ds) categor ie s	Gezondh eid_dum my	Bestaans onzekerh eid	ch13g018 Do you suffer from any kind of long- stand ing disease, afflict ion or handicap, or do you suffer from the consequ ences of an accident?
geslacht Gender	Pearson Correlation	1	-.063**	-.038	-.066**	.133**	-.019
	Sig. (2-tailed)		.002	.057	.001	<.001	.345
	N	2465	2465	2465	2465	2465	2465
leeftijd Age of the household member	Pearson Correlation	-.063**	1	-.252**	-.154**	-.096**	-.242**
	Sig. (2-tailed)	.002		<.001	<.001	<.001	<.001
	N	2465	2465	2465	2465	2465	2465
oplc at Level of education in CBS (Statistics Netherlands) categories	Pearson Correlation	-.038	-.252**	1	.131**	-.180**	.092**
	Sig. (2-tailed)	.057	<.001		<.001	<.001	<.001
	N	2465	2465	2465	2465	2465	2465
Gezondheid_dum my	Pearson Correlation	-.066**	-.154**	.131**	1	-.177**	.404**
	Sig. (2-tailed)	.001	<.001	<.001		<.001	<.001
	N	2465	2465	2465	2465	2465	2465
Bestaansonzeker heid	Pearson Correlation	.133**	-.096**	-.180**	-.177**	1	-.074**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001		<.001
	N	2465	2465	2465	2465	2465	2465
ch13g018 Do you suffer from any kind of long- standing disease, affliction or handicap, or do you suffer from the consequences of an accident?	Pearson Correlation	-.019	-.242**	.092**	.404**	-.074**	1
	Sig. (2-tailed)	.345	<.001	<.001	<.001	<.001	
	N	2465	2465	2465	2465	2465	2465

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

GEMIDDELDEN VAN ALLE VARIABELEN VERDEELD IN GEZONDE EN MINDER GEZONDE MENSEN

Als toevoeging aan de tabel met correlaties zijn de gemiddelden van alle variabelen verdeeld in gezonde en minder gezonde mensen. Hiervoor is gekozen omdat de afhankelijke variabele gezondheid een dummyvariabele is. In dit geval zeggen gemiddelden daarom meer dan correlaties.

Syntax:

```
DATASET ACTIVATE DataSet2.
SORT CASES BY Gezondheid_dummy.
SPLIT FILE SEPARATE BY Gezondheid_dummy.
```

```
FREQUENCIES VARIABLES=geslacht leeftijd oplcat Bestaansonzekerheid ch13g018
/STATISTICS=MEAN
/ORDER=ANALYSIS.
```

Statistics^a

		geslacht Gender	leeftijd Age of the household member	oplcat Level of education in CBS (Statistics Netherlands) categories	Bestaansonz ekerheid	ch13g018 Do you suffer from any kind of long- standing disease, affliction or handicap, or do you suffer from the consequence s of an accident?
N	Valid	481	481	481	481	481
	Missing	0	0	0	0	0
Mean		1.37	60.56	3.27	2.9153	1.25

a. Gezondheid_dummy = .00

geslacht Gender^a

		Frequenc y	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	305	63.4	63.4	63.4
	2 Female	176	36.6	36.6	100.0
	Total	481	100.0	100.0	

a. Gezondheid_dummy = .00

ch13g018 Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident?^a

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 yes	359	74.6	74.6	74.6
	2 no	122	25.4	25.4	100.0
	Total	481	100.0	100.0	

a. Gezondheid_dummy = .00

Statistics^a

		geslacht Gender	leeftijd Age of the household member	oplcats Level of education in CBS (Statistics Netherlands) categories	Bestaanszon- derheid	ch13g018 Do you suffer from any kind of long- standing disease, affliction or handicap, or do you suffer from the consequences of an accident?
N	Valid	1984	1984	1984	1984	1984
	Missing	0	0	0	0	0
Mean		1.29	54.70	3.78	2.5425	1.74

a. Gezondheid_dummy = 1.00

geslacht Gender^a

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	1409	71.0	71.0	71.0
	2 Female	575	29.0	29.0	100.0
	Total	1984	100.0	100.0	

a. Gezondheid_dummy = 1.00

ch13g018 Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident?^a

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 yes	513	25.9	25.9	25.9
	2 no	1471	74.1	74.1	100.0
	Total	1984	100.0	100.0	

a. Gezondheid_dummy = 1.00

VERDELING HUISHOUDHOOFDEN DOOR GESLACHT

Om meer informatie te kunnen geven over waarom er veel meer mannen in de steekproef voorkomen dan vrouwen, is bekeken of dit te maken kan hebben met dat er geselecteerd is op huishoudhoofden. Dit is zoals te zien is in de output wel het geval. Er zijn meer mannelijke huishoudhoofden dan vrouwelijke huishoudhoofden.

Syntax:

CROSSTABS

```
/TABLES=geslacht BY positie
/FORMAT=AVALUE TABLES
/CELLS=COUNT
/COUNT ROUND CELL.
```

Case Processing Summary

	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
geslacht Gender * positie Position within the household	2465	100.0%	0	0.0%	2465	100.0%

geslacht Gender * positie Position within the household Crosstabulation

Count

		positie Position within the household 1 Household head	Total
geslacht Gender	1 Male	1714	1714
	2 Female	751	751
Total		2465	2465

GESCHONDEN ASSUMPTIES VAN DE LINEAIRE REGRESSIEANLAYSE.

Uiteindelijk is er gekozen voor logistische regressieanalyse. Dit is gekozen omdat niet aan alle assumpties van lineaire regressieanalyse kon worden voldaan. De assumptie normaliteit wordt ernstig geschonden en de variabele gezondheid kent 5 categorieën, daarom is er gekozen voor een logistische regressie.

Syntax:

REGRESSION

```
/DESCRIPTIVES MEAN STDDEV CORR SIG N
```

```
/MISSING LISTWISE
```

```
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE
```

```
/CRITERIA=PIN(.05) POUT(.10)
```

```
/NOORIGIN
```

```
/DEPENDENT ch13g004
```

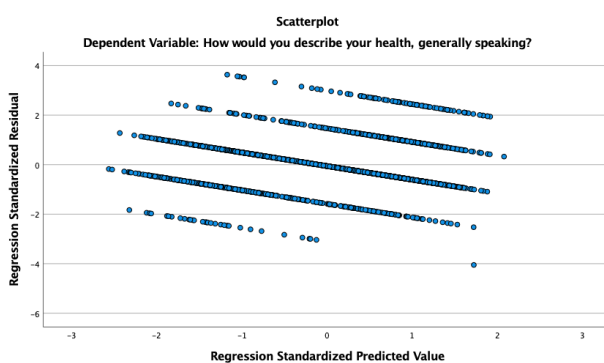
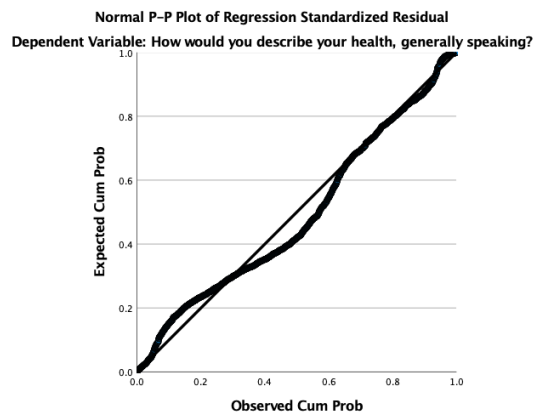
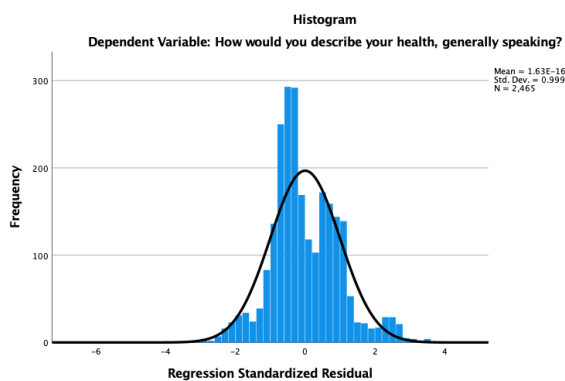
```
/METHOD=ENTER geslacht leeftijd ch13g018
```

```
/METHOD=ENTER oplcat
```

```
/METHOD=ENTER Bestaansonzekerheid
```

```
/SCATTERPLOT=(*ZRESID ,*ZPRED)
```

```
/RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID)
```



LOGISTISCHE REGRESSIEANLAYSE.

Voor dit onderzoek is gekozen voor logistische regressieanalyse, omdat de normaliteit bij de lineaire regressieanalyse ernstig geschonden was, zoals hierboven te zien is.

*De sub-paragrafen 4.2 en 4.3 bespreken de modelevaluatie en de resultaten van deze logistische analyse.

Syntax:

```

DATASET ACTIVATE DataSet1.
LOGISTIC REGRESSION VARIABLES Gezondheid_dummy
/METHOD=ENTER oplcat
/METHOD=ENTER geslacht leeftijd ch13g018
/METHOD=ENTER Bestaansonzekerheid
/SAVE=PRED PGROUP
/PRINT=GOODFIT CI(95)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
    
```

Case Processing Summary

Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	2465	100.0
	Missing Cases	0	.0
	Total	2465	100.0
Unselected Cases		0	.0
Total		2465	100.0

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

Dependent Variable Encoding

Original Value	Internal Value
.00	0
1.00	1

Block 0: Beginning Block

Classification Table^{a,b}

Observed	Gezondheid_dumm	Predicted		Percentage Correct
		.00	1.00	
Step 0 Gezondheid_dumm	.00	0	481	.0
y	1.00	0	1984	100.0
Overall Percentage				80.5

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	1.417	.051	777.341	1	<.001	4.125

Variables not in the Equation

		Score	df	Sig.
Step 0	Variables Level of education in CBS (Statistics Netherlands) categories	42.614	1	<.001
	Overall Statistics	42.614	1	<.001

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	42.380	1	<.001
	Block	42.380	1	<.001
	Model	42.380	1	<.001

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2390.965 ^a	.017	.027

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

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.296	4	.682

Contingency Table for Hosmer and Lemeshow Test

		Gezondheid_dummy = .00		Gezondheid_dummy = 1.00		Total
		Observed	Expected	Observed	Expected	
Step 1	1	66	60.245	138	143.755	204
	2	140	142.834	426	423.166	566
	3	42	44.459	166	163.541	208
	4	100	104.183	480	475.817	580
	5	95	96.694	550	548.306	645
	6	38	32.585	224	229.415	262

Classification Table^a

		Predicted Gezondheid_dumm y		Percentage Correct
Observed		.00	1.00	
Step 1	Gezondheid_dumm .00	0	481	.0
	y 1.00	0	1984	100.0
	Overall Percentage			80.5

a. The cut value is .500

		Variables in the Equation							
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	Level of education in CBS (Statistics Netherlands) categories	.216	.033	41.834	1	<.001	1.242	1.163	1.326
	Constant	.653	.124	27.709	1	<.001	1.922		

a. Variable(s) entered on step 1: Level of education in CBS (Statistics Netherlands) categories.

Block 2: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	389.181	3	<.001
	Block	389.181	3	<.001
	Model	431.562	4	<.001

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2001.784 ^a	.161	.256

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

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	9.921	8	.271

Contingency Table for Hosmer and Lemeshow Test

		Gezondheid_dummy = .00		Gezondheid_dummy = 1.00		Total
		Observed	Expected	Observed	Expected	
Step 1	1	123	127.772	124	119.228	247
	2	115	104.314	132	142.686	247
	3	84	87.393	162	158.607	246
	4	54	56.070	192	189.930	246
	5	29	26.360	219	221.640	248
	6	19	21.115	228	225.885	247
	7	16	17.874	230	228.126	246
	8	10	15.475	237	231.525	247
	9	21	13.489	224	231.511	245
	10	10	11.139	236	234.861	246

Classification Table^a

	Observed		Predicted		Percentage Correct
			Gezondheid_dum my	Gezondheid_dum 1.00	
Step 1	Gezondheid_dum	.00	66	415	13.7
	my	1.00	77	1907	96.1
Overall Percentage					80.0

a. The cut value is .500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step 1 ^a								
Level of education in CBS (Statistics Netherlands) categories	.161	.037	18.543	1	<.001	1.175	1.092	1.265
Gender	-.377	.119	10.010	1	.002	.686	.543	.866
Age of the household member	-.010	.004	6.230	1	.013	.990	.982	.998
Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident?	2.049	.120	293.642	1	<.001	7.764	6.142	9.815
Constant	-1.123	.410	7.494	1	.006	.325		

a. Variable(s) entered on step 1: Gender, Age of the household member, Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident? .

Block 3: Method = Enter**Omnibus Tests of Model Coefficients**

	Chi-square	df	Sig.
Step 1 Step	56.065	1	<.001
Block	56.065	1	<.001
Model	487.627	5	<.001

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1945.718 ^a	.179	.286

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

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.811	8	.946

Contingency Table for Hosmer and Lemeshow Test

		Gezondheid_dummy = .00		Gezondheid_dummy = 1.00		Total
		Observed	Expected	Observed	Expected	
Step 1	1	144	141.600	103	105.400	247
	2	104	105.264	143	141.736	247
	3	80	82.022	167	164.978	247
	4	53	51.852	194	195.148	247
	5	33	29.567	214	217.433	247
	6	19	21.475	228	225.525	247
	7	16	17.254	232	230.746	248
	8	11	13.960	236	233.040	247
	9	14	10.729	233	236.271	247
	10	7	7.277	234	233.723	241

Classification Table^a

Observed	Gezondheid_dumm	Predicted		Percentage Correct
		Gezondheid_dumm y	1.00	
Step 1	Gezondheid_dumm .00	131	350	27.2
	y 1.00	89	1895	95.5
	Overall Percentage			82.2

a. The cut value is .500

		Variables in the Equation							
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	Level of education in CBS (Statistics Netherlands) categories	.106	.039	7.516	1	.006	1.112	1.031	1.199
	Gender	-.283	.122	5.426	1	.020	.753	.594	.956
	Age of the household member	-.018	.004	16.301	1	<.001	.983	.974	.991
	Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident?	2.015	.121	278.240	1	<.001	7.497	5.917	9.500
	Bestaansonzekerheid	-.518	.070	54.693	1	<.001	.596	.519	.683
	Constant	.830	.489	2.877	1	.090	2.294		

a. Variable(s) entered on step 1: Bestaansonzekerheid.

LINEARE REGRESSIE MODEL 3

Model 3 is uitgevoerd doormiddel van lineaire regressie, omdat dit model een andere afhankelijke variabele heeft, namelijk bestaansonzekerheid.

Syntax:

```
REGRESSION
/DESCRIPTIVES MEAN STDDEV CORR SIG N
/MISSING LISTWISE
/STATISTICS COEFF OUTS CI(95) R ANOVA CHANGE
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT Bestaansonzekerheid
/METHOD=ENTER oplcat leeftijd geslacht ch13g018.
```

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	oplcac Level of education in CBS (Statistics Netherlands) categories, geslacht Gender, ch13g018 Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident? , leeftijd Age of the household member ^b		Enter

a. Dependent Variable: Bestaansonzekerheid

b. All requested variables entered.

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	.274 ^a	.075	.074	.80347	.075	49.986	4	2460	<.001

a. Predictors: (Constant), ch13g018 Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident? , geslacht Gender, oplcac Level of education in CBS (Statistics Netherlands) categories, leeftijd Age of the household member

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	129.078	4	32.270	49.986	<.001 ^b
	Residual	1588.107	2460	.646		
	Total	1717.186	2464			

a. Dependent Variable: Bestaansonzekerheid

b. Predictors: (Constant), ch13g018 Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident? , geslacht Gender, oplcac Level of education in CBS (Statistics Netherlands) categories, leeftijd Age of the household member

Model		Coefficients ^a					95.0% Confidence Interval for B	
		Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	Lower Bound	Upper Bound
		B	Std. Error					
1	(Constant)	3.540	.123		28.889	<.001	3.300	3.780
	oplcatt Level of education in CBS (Statistics Netherlands) categories	-.114	.011	-.209	-10.408	<.001	-.136	-.093
	leeftijd Age of the household member	-.009	.001	-.164	-7.943	<.001	-.011	-.007
	geslacht Gender	.204	.035	.113	5.785	<.001	.135	.273
	ch13g018 Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident?	-.161	.035	-.092	-4.599	<.001	-.229	-.092

a. Dependent Variable: Bestaansonzekerheid

BIJLAGE 3

CONTROLE OP OUTLIERS

Voor de controle op outliers zijn meerdere manieren gebruikte: leverage, gestandaardiseerde residuen en DFBETA.

Syntax.

```
LOGISTIC REGRESSION VARIABLES Gezondheid_dummy  
/METHOD=ENTER oplcat  
/METHOD=ENTER geslacht ch13g018 leeftijd  
/METHOD=ENTER Bestaansonzekerheid  
/SAVE=PRED COOK LEVER DFBETA  
/PRINT=GOODFIT  
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

```
LOGISTIC REGRESSION VARIABLES Gezondheid_dummy  
/METHOD=ENTER oplcat  
/METHOD=ENTER geslacht leeftijd ch13g018  
/METHOD=ENTER Bestaansonzekerheid  
/SAVE=PRED PGROUP RESID ZRESID  
/CASEWISE OUTLIER(2)  
/PRINT=GOODFIT CI(95)  
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

Leverage.

Leverage manier geeft de invloed van een punt weer op de regressielijn. De leverage waarden moet liggen tussen 0 en $2 \cdot p/n = 6/2465 = 0.0024$. Meer dan 40,9% heeft een hogere leverage waarde dan 0,0024, zoals in de output zichtbaar is gaat het om 1004 cases. Als deze cases afzonderlijk worden bekeken lijken ze niet heel opvallend en lijken er geen hele schokkende dingen zichtbaar. Ook is zijn er geen hele schokkende waardes die ver van de andere waardes afstaan, de verschillen zijn steeds heel minimaal met de waarde ervoor. Ook lijkt het voor de onderzoeksvraag geen toegevoegde waarde te hebben om de cases te verwijderen. Daarom is ervoor gekozen deze cases niet te verwijderen uit de dataset.

Syntax:

```
FREQUENCIES VARIABLES=LEV_3  
/STATISTICS=MINIMUM MAXIMUM  
/ORDER=ANALYSIS.
```

```
DATASET ACTIVATE DataSet1.  
USE ALL.  
COMPUTE filter_$(LEV_3 >= 0.0024).  
VARIABLE LABELS filter_ $ 'LEV_3 >= 0.0024 (FILTER)'.  
VALUE LABELS filter_ $ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_ $(f1.0).  
FILTER BY filter_$.  
EXECUTE.
```

```
FREQUENCIES VARIABLES=LEV_3  
/ORDER=ANALYSIS.
```

Statistics		
LEV_4 Leverage value		
N	Valid	2465
	Missing	0
Minimum		.00065
Maximum		.01108

Statistics		
LEV_3 Leverage value		
N	Valid	1004
	Missing	0

Gestandaardiseerde residuen

De gestandaardiseerde residuen mogen niet groter zijn dan 2 en niet kleiner dan -2. Door te kijken naar de waarden van de gestandaardiseerde residuen, valt te concluderen dat er 120 cases zijn met gestandaardiseerde residuen kleiner dan -2. Hiervoor geldt eigenlijk hetzelfde te zeggen als voor de leverage. Als deze cases afzonderlijk worden bekeken lijken ze niet heel opvallend en lijken er geen hele schokkende dingen zichtbaar. Ook is zijn er geen hele schokkende waarden die ver van de andere waarden afstaan, de verschillen zijn steeds heel minimaal met de waarde ervoor. Ook lijkt het voor de onderzoeksvraag geen toegevoegde waarde te hebben om de cases te verwijderen. Daarom is ervoor gekozen deze cases niet te verwijderen uit de dataset.

Syntax:

```
FREQUENCIES VARIABLES=ZRE_3
/STATISTICS=MINIMUM MAXIMUM
/ORDER=ANALYSIS.
```

Statistics		
ZRE_3 Normalized residual		
N	Valid	2465
	Missing	0
Minimum		-6.93851
Maximum		1.63211

DFBETA

De DFBETA geeft het effect op de helling weer als observatie i wordt weggehaald. Hiervoor is geen criterium, er moet worden gekeken naar de grootste waarde. De grootste waarde is 0.037, dit kan wel een grote invloed hebben op de data. Verder liggen alle punten dicht bij elkaar en lijkt er geen sprake van outliers. Daarom zijn er volgens de DFBETA geen mogelijke outliers.

Syntax:

```
FREQUENCIES VARIABLES=DFB0_4 DFB1_4 DFB2_4 DFB3_4 DFB4_4 DFB5_4
/STATISTICS=MINIMUM MAXIMUM MEAN
/ORDER=ANALYSIS.
```

Statistics

		DFB0_4 DFBETA for constant	DFB1_4 DFBETA for Level of education in CBS (Statistics Netherlands) categories	DFB2_4 DFBETA for Gender	DFB3_4 DFBETA for Do you suffer from any kind of long- standing disease, affliction or handicap, or do you suffer from the consequenc es of an accident?	DFB4_4 DFBETA for Age of the household member	DFB5_4 DFBETA for Bestaanson zekerheid
N	Valid	2465	2465	2465	2465	2465	2465
	Missin g	0	0	0	0	0	0
Mean		.0000008	.0000000	-.0000001	.0000000	.0000000	-.0000001
Minimum		-.06539	-.00408	-.01052	-.01082	-.00042	-.00945
Maximum		.05045	.00465	.00694	.00649	.00064	.00774

Waarden die zowel een uitbijter zijn bij leverage als gestandaardiseerde resdiuen.

Pas als cases op meerdere manieren afwijkende waardes hebben kunnen ze gezien worden als outliers. In deze dataset zijn dit er 34. Wanneer afzonderlijk naar deze cases gekeken wordt lijken er geen echte uitbijters bij te zijn met hele opvallende scores. Dat er toch afwijkende waardes zijn kan te maken hebben met de verdeling van gezondheid, waarbij veel meer mensen een goede gezondheid hebben dan een minder goede gezondheid. Met deze kennis en het feit dat er geen hele opvallende waardes te vinden zijn, is besloten om de afwijkende waardes niet als outliers te beschouwen. Ook omdat de verdeling van gezondheid dan nog schever zou worden.

Syntax:

```
USE ALL.  
COMPUTE filter_$(ZRE_4 <= -2 & LEV_3 >= 0.0024).  
VARIABLE LABELS filter_$(ZRE_4 <= -2 & LEV_3 >= 0.0024 (FILTER)).  
VALUE LABELS filter_$(0 'Not Selected' 1 'Selected').  
FORMATS filter_$(f1.0).  
FILTER BY filter_$.  
EXECUTE.
```

```
FREQUENCIES VARIABLES=LEV_3  
/ORDER=ANALYSIS.
```

Statistics

LEV_3 Leverage value

N	Valid	34
	Missing	0

CONTROLE OP MULTICOLINEARITEIT

Doormiddel van de VIF-scores wordt er gecontroleerd voor multicolineariteit. Hier lijkt geen sprake van, er is geen opvallend hoge VIF-score. Geen enkele VIF-score is namelijk hoger dan 4.

```
DATASET ACTIVATE DataSet1.  
REGRESSION  
/MISSING LISTWISE  
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL  
/CRITERIA=PIN(.05) POUT(.10)  
/NOORIGIN  
/DEPENDENT Gezondheid_dummy  
/METHOD=ENTER oplcat  
/METHOD=ENTER geslacht leeftijd ch13g018  
/METHOD=ENTER Bestaansonzekerheid.
```


Model		Coefficients ^a					Collinearity Statistics	
		Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	Tolerance	VIF
		B	Std. Error					
1	(Constant)	.679	.021		32.880	<.001		
	oplcac Level of education in CBS (Statistics Netherlands) categories	.034	.005	.131	6.582	<.001	1.000	1.000
2	(Constant)	.328	.055		5.985	<.001		
	oplcac Level of education in CBS (Statistics Netherlands) categories	.022	.005	.083	4.384	<.001	.932	1.072
	geslacht Gender	-.050	.016	-.058	-3.143	.002	.992	1.008
	leeftijd Age of the household member	-.001	.001	-.043	-2.227	.026	.882	1.133
	ch13g018 Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident?	.319	.016	.385	20.410	<.001	.940	1.064
3	(Constant)	.565	.063		8.994	<.001		
	oplcac Level of education in CBS (Statistics Netherlands) categories	.014	.005	.054	2.802	.005	.893	1.120
	geslacht Gender	-.036	.016	-.042	-2.290	.022	.978	1.022
	leeftijd Age of the household member	-.002	.001	-.066	-3.406	<.001	.860	1.162
	ch13g018 Do you suffer from any kind of long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident?	.308	.016	.372	19.858	<.001	.932	1.073
	Bestaansonzekerheid	-.067	.009	-.141	-7.481	<.001	.925	1.081

a. Dependent Variable: Gezondheid_dummy

