



university of  
groningen

faculty of behavioural  
and social sciences

# GEVOELEN VAN BOOSHEID OMTRENT DE STIJGING VAN DE AOW-LEEFTIJD

*Wat is de invloed van opleidingsniveau op de mate van boosheid ten aanzien van de stijging van de AOW-leeftijd? En in hoeverre wordt deze invloed verklaard door het inkomen en de mate van fysiek zwaar werk van de respondent?*

**Naam:** Esmee Boomsma (S3815641)

**E-mailadres:** e.boomsma@student.rug.nl

**Begeleider:** Gijs Huitsing

**Tweede lezer:** Gert Stulp

**Studie:** Sociologie

**Vak:** Bachelorwerkstuk

**Datum:** 08-06-2022

## Samenvatting

De AOW-leeftijd is de afgelopen tien jaar stapsgewijs toegenomen en het is de verwachting dat het zal blijven stijgen. Als reactie op deze ontwikkeling zijn negatieve emoties te verwachten. Deze studie zal dieper ingaan op de frustratie die heerst rondom de stijging en er is gekeken onder wie deze frustratie het meeste heerst. Zo is er specifiek gekeken naar het opleidingsniveau, het inkomen en de mate waarin mensen hun werk als fysiek veeleisend ervaren. Voor het onderzoek is de volgende onderzoeksvraag opgesteld: *Wat is de invloed van opleidingsniveau op de mate van boosheid t.a.v. de stijging van de AOW-leeftijd? En in hoeverre wordt deze invloed verklaard door het inkomen en de mate van fysiek zwaar werk?* Om antwoord te kunnen geven op de onderzoeksvraag is gebruik gemaakt van online vragenlijsten van het LISS (Longitudinal Internet studies for the Social Sciences) panel. Uit de resultaten is naar voren gekomen dat hogeropgeleiden minder boos zijn over de stijging van de AOW-leeftijd dan lageropgeleiden. Verder is er geen bewijs gevonden dat inkomen en de mate van fysiek zwaar werk een deel van de invloed van opleidingsniveau op boosheid verklaren. Dit kan liggen aan een aantal beperkingen van het onderzoek, zoals de schending van modelassumpties. Het onderzoek kan als basis worden gezien voor het verder onderzoeken van de gevoelens van boosheid omtrent de stijging van de AOW-leeftijd.

## Inhoudsopgave

1. Inleiding .....	3
2. Theoretisch kader .....	5
2.1 Invloed van opleidingsniveau op de mate van boosheid .....	5
2.2 Het inkomen .....	6
2.3 De mate van fysiek zwaar werk .....	7
2.4 Controlevariabelen .....	8
3. Methoden .....	9
3.1 Beschrijving van de data .....	9
3.2 Beschrijving van de vragenlijst .....	9
3.3 Operationalisaties .....	10
3.4 Analyse-opzet .....	12
4. Resultaten .....	13
4.1 Beschrijvende statistieken .....	13
4.2 Modevaluatie .....	15
4.3 Hypothesetoetsing .....	19
5. Conclusie en discussie .....	21
Literatuur .....	24
Bijlage 1 – Overzicht variabelen .....	26
Afhankelijke variabele: Boosheid .....	27
Onafhankelijke variabele: Opleidingsniveau .....	29
Moderator: Inkomen .....	31
Moderator: Fysiek zwaar werk .....	33
Controlevariabele: Geslacht .....	35
Controlevariabele: Leeftijd .....	36
Controlevariabele: Werksector .....	39
Bijlage 2 – Statistische analyses .....	42
Onderdeel 1: Bivariate statistieken .....	42
Onderdeel 2: Modelschattingen .....	67
Bijlage 3 – Multicollineariteit, modelassumpties en uitbijters .....	76
Multicollineariteit .....	76
Modelassumpties .....	77
Uitbijters .....	79

## 1. Inleiding

Sinds 1957 is de Algemene Ouderdomswet-leeftijd 65 jaar en is 55 jaar lang niet gewijzigd (De Beer & Van der Gaag, 2019). Gezien de levensverwachting de afgelopen decennia fors is toegenomen, onder andere door de sterk verbeterde kwaliteit van de gezondheidszorg (Stam, Garssen, Kardal & Lodder, 2008), is het een logisch gevolg om de AOW-leeftijd te laten stijgen. Men kan namelijk langer doorwerken doordat ze ouder worden. Een andere reden om de AOW-leeftijd te laten stijgen is dat het aantal ouderen toeneemt in de samenleving, terwijl de beroepsbevolking nauwelijks groeit (De Beer & Van der Gaag, 2019). Nederlanders zullen zo langer moeten doorwerken om de kosten voor de AOW en pensioenen betaalbaar te houden (Hoeymans, 2011).

Sinds 2012 is de AOW-leeftijd dan ook stapsgewijs verhoogd tot 67 jaar in 2024 en zal deze vanaf 2024 worden gekoppeld aan ontwikkelingen in de levensverwachting op leeftijd 65 (De Beer & Van der Gaag, 2019). De stijging van de AOW-leeftijd is echter vaak een discussiepunt geweest in de politiek en in de samenleving. Er heerst namelijk veel onvrede rondom de stijging van de AOW-leeftijd (De Beer & Van der Gaag, 2019). Zo is het bijvoorbeeld niet voor iedereen mogelijk om langer door te werken. Dit geldt bijvoorbeeld voor werknemers met gezondheidsklachten of met een zwaar beroep (Visser, 2019). De onvrede kan ook komen door oudere werknemers die een fysieke belasting hebben door het werk. Voor hen wordt het (lichamelijk) moeilijk om langer door te moeten werken (Visser, 2019).

Dit onderzoek beschrijft de onvrede die heerst onder de stijging van de AOW-leeftijd. Er zal worden gekeken naar wie er boos zijn over de stijging van de AOW-leeftijd. Uit de eerste alinea blijkt dat het noodzakelijk is dat men langer doorwerkt in verband met de vergrijzing en de kosten die hieraan verbonden zijn. Het is echter lastig voor een bepaalde groep mensen om langer te moeten werken, waardoor men zorgen kan ervaren. Ook kan het zorgen voor frustratie onder deze groep of er kan bij deze groep zelfs een afnemend vertrouwen in de overheid ontstaan. Het is daarom belangrijk dat er onder de aandacht komt welke groepen er precies boos zijn en dat hier wellicht iets aan kan worden gedaan, zodat eventuele frustratie en afnemend vertrouwen worden weggenomen.

Er is gekozen om het onderzoek te richten op het opleidingsniveau, een belangrijke indicator van de sociaaleconomische status. Hierbij draait het om de invloed van opleidingsniveau op de mate van boosheid ten aanzien van de stijging van de AOW-leeftijd. Er wordt ook gekeken naar inkomen en de fysieke zwaarte van het werk. Hierbij wordt gekeken in hoeverre zij de invloed van opleidingsniveau op de mate van boosheid mediëren.

Er is al onderzoek gedaan naar de mate van onvrede / boosheid over de stijging van de AOW-leeftijd. Hierbij is er ook gekeken naar welke invloed opleidingsniveau heeft op de mate van boosheid. Het blijkt dat vooral laag- en middelbaar opgeleide werknemers frustratie ervaren ten

aanzien van de stijging van de AOW-leeftijd. Dit is in tegenstelling tot universitair geschoolde werknemers, die veel minder negatieve emoties ervaren ten aanzien van de stijging (De Beer, Van Dalen & Henkens, 2017a). In mijn onderzoek wordt er op bovenstaand onderzoek voortgeborduurd. Er wordt specifiek gekeken naar het inkomen en de fysieke zwaarte van het werk en in hoeverre zij de invloed van opleidingsniveau op de mate van boosheid verklaren. Door ook naar het inkomen en de fysieke zwaarte van het werk te kijken, wordt er meer diepte gegeven aan het verband tussen opleidingsniveau en de mate van boosheid. Daarnaast is het onderzoek van De Beer en collega's (2017a) in 2015 uitgevoerd. We zijn inmiddels een aantal jaren verder en de AOW-leeftijd is alweer aangepast. Ook is mijn onderzoek gericht op een leeftijdscategorie van 40-69 jaar en het onderzoek van De Beer en collega's (2017a) op werknemers van 60-65 jaar. Hierdoor kan er in mijn onderzoek ook iets worden gezegd over veertigers en vijftigers en de mate waarin zij boos zijn over de stijging van de AOW-leeftijd.

Het doel van dit onderzoek is dus om te kijken welke invloed het opleidingsniveau heeft op de mate van boosheid t.a.v. de stijging van de AOW-leeftijd en wat de rol van fysiek zwaar werk en inkomen is. Hierbij is het belangrijk dat er onder de aandacht komt wie er precies boos zijn en wat de redenen hiervoor zijn. De probleemstelling luidt dan ook als volgt: *Wat is de invloed van opleidingsniveau op de mate van boosheid t.a.v. de stijging van de AOW-leeftijd? En in hoeverre wordt deze invloed verklaard door het inkomen en de mate van fysiek zwaar werk van de respondent?*

## 2. Theoretisch kader

Dit hoofdstuk geeft het theoretisch kader van het onderzoek weer. Allereerst volgt uitleg betreffende de invloed van opleidingsniveau op de mate van boosheid. Dit wordt toegelicht aan de hand van het aantal dienstjaren van het individu, de (gezonde)levensverwachting en de mate van autonomie in het werk. Vervolgens volgt de toelichting van het inkomen en de mate van fysiek zwaar werk. Tot slot volgt een korte toelichting van de controlevariabelen.

### 2.1 Invloed van opleidingsniveau op de mate van boosheid

Allereerst hebben lageropgeleiden over het algemeen meer dienstjaren dan hogeropgeleiden (Henkens, Van Solinge, Damman & Dingemans, 2016). Lageropgeleiden beginnen gemiddeld gezien vaak al eerder met werken dan hogeropgeleiden (De Beer & Van der Gaag, 2019). Hogeropgeleiden zijn over het algemeen langer bezig met hun studie en stappen dan ook later het arbeidsleven in. Wanneer de AOW-leeftijd stijgt, zal men nog langer moeten doorwerken. Lageropgeleiden zijn vroeg begonnen met werken en zullen daardoor in verhouding tot hogeropgeleiden meer jaren van hun leven werken.

Ten tweede speelt ook de levensverwachting een rol. De AOW-leeftijd is gebaseerd op de gemiddelde levensverwachting van de samenleving. De levensverwachting is afgelopen decennia behoorlijk toegenomen onder de bevolking. Er zijn echter grote verschillen te zien tussen hoog- en laagopgeleiden (Bruggink, Garssen, Lodder & Kardal, 2009). De gemiddelde levensverwachting van lageropgeleiden ligt namelijk een stuk lager dan die van hogeropgeleiden. De levensverwachting van lageropgeleiden ligt zo'n zeven jaar lager dan die van hogeropgeleiden (Bruggink, 2009). Dit zorgt ervoor dat lageropgeleiden minder AOW zullen ontvangen (Smulders, Houtman & Van den Bossche, 2009). De AOW-leeftijd is voor beide groepen hetzelfde, maar de lageropgeleiden kunnen gemiddeld minder lang van hun pensioen genieten door de lagere levensverwachting.

Ook speelt de gezondheid een rol. Naarmate men ouder wordt, krijgt men meer te maken met gezondheidsproblemen (De Beer & Van der Gaag, 2019). Dit houdt in dat wanneer de AOW-leeftijd hoger wordt, het lastiger wordt om langer door te moeten werken. Ook neemt het risico op arbeidsongeschiktheid toe naarmate men ouder wordt (De Beer & Van der Gaag, 2019). Dit geldt met name voor lageropgeleiden. Lageropgeleiden hebben namelijk een lagere gezonde levensverwachting dan hogeropgeleiden (Bruggink, 2009).

De gezonde levensverwachting is het aantal jaren dat een persoon van een bepaalde leeftijd naar verwachting nog in goede gezondheid zal leven. De gezonde levensverwachting wordt gemeten aan de hand van drie maten. Namelijk de levensverwachting in als goed ervaren gezondheid, een levensverwachting zonder chronische ziektes en een levensverwachting zonder lichamelijke

beperkingen (Bruggink, 2009). Hogeropgeleiden brengen ongeveer 70% van hun AOW-jaren door in goede gezondheid, terwijl lageropgeleiden zo'n 45% in goede gezondheid hun AOW-jaren doorbrengen (De Beer, Van Dalen & Henkens, 2017b).

Door de stijging van de AOW-leeftijd kunnen er relatief minder jaren in goede gezondheid worden doorgebracht (De Beer & Van der Gaag, 2019). Dit geldt met name voor lageropgeleiden, blijkt uit bovenstaande alinea. Ze kunnen hierdoor ook minder genieten van hun pensioen. Wanneer men met pensioen gaat heeft men meer vrije tijd en kan bijvoorbeeld leuke dingen ondernemen. Hierin worden echter met name lageropgeleiden in tegengehouden, doordat ze een grotere kans hebben op lichamelijke beperkingen en chronische ziektes naarmate ze ouder worden. Ook leidt een slechtere gezondheid ertoe dat het werk als meer belastend kan worden ervaren, waardoor er ook behoefte aan meer vrije tijd, en dus eerder uittreden, kan heersen onder lageropgeleiden (Kok et al., 2017).

Tot slot kan de mate van autonomie wat wordt ervaren in het werk een rol spelen. Het ervaren van autonomie in een baan houdt onder andere in dat men het eigen werktempo kan bepalen en het werk kan inrichten hoe men het wilt (Mol & De Vries, 2009). Dit leidt er onder andere toe dat het werk als minder belastend kan worden ervaren. Hogeropgeleiden ervaren over het algemeen meer autonomie in hun baan dan lageropgeleiden (Mol & De Vries, 2009). Dit kan een reden zijn dat lageropgeleiden het werk als meer belastend kunnen ervaren.

Uit bovenstaande tekst is gebleken dat hogeropgeleiden gemiddeld genomen minder dienstjaren hebben en een hogere (gezonde) levensverwachting hebben ten opzichte van lageropgeleiden. Ook ervaren hogeropgeleiden meer autonomie in hun baan dan lageropgeleiden. Dit leidt ertoe dat het voor hogeropgeleiden minder belastend kan zijn om langer door te moeten werken dan voor lageropgeleiden. Hierdoor kan er onder hogeropgeleiden minder boosheid heersen over de stijging van de AOW-leeftijd dan onder lageropgeleiden. De eerste hypothese luidt dan ook als volgt: *Hoe hoger het opleidingsniveau van de respondent, hoe minder boos men is over de stijging van de AOW-leeftijd.*

## **2.2 Het inkomen**

Lageropgeleiden verdienen over het algemeen minder dan hoogopgeleiden. Mensen met een lager inkomen hebben een grotere voorkeur om eerder te stoppen met werken dan mensen met een hoger inkomen (Kok et al., 2017). Onder andere door de lagere levensverwachting en de behoefte aan vrije tijd (Kok et al., 2017). Doordat lageropgeleiden minder verdienen, kunnen ze ook minder snel stoppen met werken. Ze beschikken niet over de financiële middelen om eerder te kunnen stoppen (De Beer & Van der Gaag, 2019). Zo hebben ze vaker onvoldoende pensioen opgebouwd zodat men eerder kan stoppen met het werk wanneer dit wenselijk is (Kok et al., 2017).

Lageropgeleiden hebben dus meer de wens om eerder te stoppen met werken dan hogeropgeleiden, maar worden hierin tegengehouden door wat ze verdienen. Door de stijging van de AOW-leeftijd, zullen ze nog langer moeten doorwerken. Hogeropgeleiden hebben meer de financiële vrijheid dan lageropgeleiden om eerder te stoppen met het werk wanneer dit gewenst is. Dit kan ervoor zorgen dat hogeropgeleiden minder boosheid ervaren voor de stijging van de AOW-leeftijd. De tweede hypothese luidt dan ook als volgt: *Het negatieve effect van het hebben van een hoger opleidingsniveau op de mate van boosheid t.a.v. de stijging van de AOW-leeftijd, wordt (gedeeltelijk) verklaard door inkomen.*

### **2.3 De mate van fysiek zwaar werk**

Lageropgeleiden hebben over het algemeen meer fysiek zwaar werk dan hogeropgeleiden (Kok et al., 2017). Ook kan het werk gevaarlijk zijn. Een voorbeeld van dit soort werk is werk in de bouw of werken met gevaarlijke stoffen.

Fysiek zwaar werk vraagt veel van het lichaam. Al helemaal wanneer men ouder wordt. Wanneer men ouder wordt neemt de kans op gezondheidsproblemen namelijk toe (De Beer & Van der Gaag, 2019). Hoe ouder men wordt, hoe meer men te maken krijgt met lichamelijke mankementen. Aangezien fysiek zwaar werk veel van het lichaam vraagt, kan het zijn dat lageropgeleiden het lichamelijke (zware) werk nog maar net kunnen uitvoeren op oudere leeftijd. De werkcapaciteit daalt met de leeftijd. Daarnaast is er gemiddeld gezien onder lageropgeleiden verminderde sprake van *employability* (Smit, Andriessen & Stark, 2005). Dit houdt kort gezegd in dat men breed inzetbaar is op de arbeidsmarkt en dit ook blijft, door middel van scholing bijvoorbeeld. Lageropgeleiden hebben over het algemeen minder de ambitie en motivatie om zich om te laten scholen in tegenstelling tot hogeropgeleiden (Smit, Andriessen & Stark, 2005), waardoor ze bijvoorbeeld een kleinere kans hebben om andersoortig werk uit te oefenen op latere leeftijd.

Daarbij is er voor mensen die fysiek zwaar werk uitvoeren ook de wens om eerder te stoppen (Damman & Oude Mulders, 2019). Dit is echter vaak niet haalbaar. Mensen die fysiek zwaar werk uitoefenen zijn immers vaker laagopgeleid. De verwachte pensioenen en lonen zijn niet hoog genoeg zodat het ook kan worden bewerkstelligd dat ze eerder kunnen stoppen (Damman & Oude Mulders, 2019).

Lageropgeleiden hebben dus over het algemeen meer fysiek zware beroepen dan hogeropgeleiden. Het wordt voor deze groep moeilijker om op oudere leeftijd het werk nog uit te voeren en daarnaast hebben ze vaak niet de scholing om ander soortig werk te gaan doen. Het kan dan ook zorgen voor boosheid wanneer deze groep langer door moet werken, terwijl dit lichamenlijk nog maar net lukt en er ook geen opties zijn betreffend ander soort werk. De derde hypothese luidt



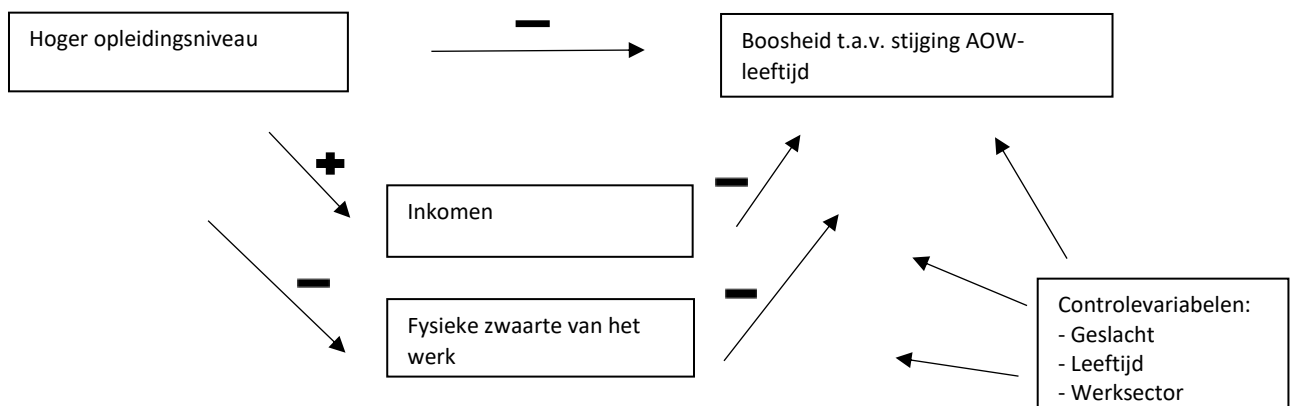
dan ook als volgt: *Het negatieve effect van het hebben van een hoger opleidingsniveau op de mate van boosheid t.a.v. de stijging van de AOW-leeftijd, wordt (gedeeltelijk) verklaard door de mate van fysiek zwaar werk.*

## 2.4 Controlevariabelen

De eerste controlevariabele is geslacht. Hierbij is gecontroleerd voor een hoofdeffect van geslacht op de mate van boosheid over de stijging van de AOW-leeftijd.

De tweede controlevariabele is leeftijd. Hierbij is gecontroleerd voor een hoofdeffect van leeftijd op de mate van boosheid over de stijging van de AOW-leeftijd. Het is belangrijk om hiervoor te controleren, omdat er een verschil kan zitten in de verschillende leeftijden en de mate van boosheid. De mate van stijging van de AOW-leeftijd hangt namelijk af van de geboortedatum. Zo hebben jongere werknemers een hogere AOW-leeftijd dan oudere werknemers.

De derde controlevariabele is werksector. Hierbij is gecontroleerd voor het hoofdeffect van werksector op de mate van boosheid betreffende de stijging van de AOW-leeftijd. Men kan per werksector verschillen in de mate waarop ze boos zijn om langer door te werken. In figuur 1 staat het conceptuele model van het onderzoek weergegeven.



Figuur 1: Onderzoeksmodel

### 3. Methoden

Dit hoofdstuk geeft aan welke methoden zijn gebruikt om tot de resultaten te komen. Als eerst volgt een beschrijving van de dataset en een beschrijving van de vragenlijst. Vervolgens volgen de operationalisaties en tot slot volgt de analyse-opzet van het onderzoek.

#### 3.1 Beschrijving van de data

Het LISS (Longitudinal Internet studies for the Social Sciences) panel onderzoekt verschillende thema's, waaronder gezondheid, politiek, religie, scholing, werk en pensioenen. De gehele dataset, die wordt gebruikt in dit onderzoek, bestaat uit zo'n 5000 gezinnen en omvat ongeveer 7500 individuen. De huishoudens worden willekeurig geselecteerd, maar er kunnen meerdere individuen uit een huishouden meedoen aan een panel (LISS, z.d.). Elk jaar is het onderzoek herhaald voor dezelfde respondenten, zodat er uiteindelijk een verandering in de antwoorden van de respondenten kan worden gemeten. De vragenlijsten zijn over meerdere jaren bevestigd, waardoor er meerdere Waves beschikbaar zijn. In dit onderzoek is gebruik gemaakt van Wave 13, wat in 2020 is uitgevoerd. Het LISS panel onderzoekt dus vele onderwerpen, waarvoor verschillende vragenlijsten zijn opgesteld. Voor mijn onderzoek is alleen gebruik gemaakt van de vragenlijst die betrekking heeft op 'werk en scholing' en op de vragenlijst die betrekking heeft op 'langer doorwerken (betreffende de stijging van de AOW-leeftijd) en pensioenplanning'.

De populatie van dit onderzoek bestaat uit werkende inwoners van Nederland in de leeftijdscategorie van 40-69 jaar. De respondenten zijn willekeurig geselecteerd voor de vragenlijst. De vragenlijst 'langer doorwerken en pensioenen' is niet aan alle 7500 respondenten bevestigd. De vragenlijst is bevestigd aan het deel van de respondenten die betaald werk hebben. Hierdoor zitten er een stuk minder respondenten in de steekproef, uiteindelijk 1259 respondenten. Het zijn nog genoeg respondenten om een representatieve steekproef neer te zetten.

De respondenten zijn geselecteerd door middel van een aselechte steekproef, op basis van data van het Centraal Bureau voor de Statistiek. Een aselechte steekproef houdt in dat ieder individu in de populatie dezelfde kans heeft om in de steekproef terecht te komen.

#### 3.2 Beschrijving van de vragenlijst

De vragenlijst is aan de respondenten voorgelegd en ze konden de vragenlijst online invullen. De duur van de vragenlijst is zo'n 15-30 minuten. De respondenten krijgen betaald voor het invullen van een complete vragenlijst. De financiële vergoeding bedraagt 15 euro per uur. Het bedrag is gebaseerd op de geschatte tijd die een gemiddeld panellid nodig heeft om de vragenlijst in te vullen (LISS, z.d.). De vragen bestaan uit onder andere de achtergrondkenmerken van de respondent, zoals

leeftijd, geslacht, inkomen en opleidingsniveau. Daarnaast heeft de respondent ook vragen ingevuld over het langer moeten doorwerken en over pensioenen. Deze vragen hebben betrekking op hoe men zich erover voelt dat ze langer moeten doorwerken (door de stijging van de AOW-leeftijd) en over de kennis die ze hebben betreffende de AOW-leeftijd.

### 3.3 Operationalisaties

De afhankelijke variabele is de mate van boosheid ten aanzien van de stijging van de AOW-leeftijd en de onafhankelijke variabele is het opleidingsniveau van de respondent. Er is gecontroleerd voor geslacht, leeftijd en werksector.

#### Afhankelijke variabele: boosheid

De afhankelijke variabele geeft de boosheid ten aanzien van de stijging van de AOW-leeftijd weer. Hierbij wordt gevraagd: 'Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lang moet doorwerken?' De variabele heeft de volgende antwoordcategorieën: 1 = "niet", 2 = "in geringe mate", 3 = "in redelijke mate", 4 = "in hoge mate".

#### Onafhankelijke variabele: opleidingsniveau

De onafhankelijke variabele geeft het opleidingsniveau van de respondent aan. Bij deze variabele is het label 'Opleiding in CBS-categorieën'. De variabele heeft de volgende antwoordcategorieën: 1 = "basisonderwijs", 2 = "vmbo", 3 = "havo/vwo", 4 = "mbo", 5 = "hbo", 6 = "wo", 9 = "Onbekend (missing)".

#### Mediator: inkomen

De eerste mediator geeft het inkomen van de respondent aan met het label 'Persoonlijk netto maandinkomen'. De variabele heeft de volgende antwoordcategorieën: 0 = "geen inkomen", 1 = "verdient 500 euro of minder", 2 = "verdient 501 t/m 1000 euro", 3 = "verdient 1001 t/m 1500 euro", 4 = "verdient 1501 t/m 2000 euro", 5 = "verdient 2001 t/m 2500 euro", 6 = "verdient 2501 t/m 3000 euro", 7 = "verdient 3001 t/m 3500 euro", 8 = "verdient 3501 t/m 4000 euro", 9 = "verdient 4001 t/m 4500 euro", 10 = "verdient 4501 t/m 5000 euro", 11 = "verdient 5001 t/m 7500 euro", 12 = "Meer dan 7500 euro", 13 = "Dat weet ik echt niet", 14 = "Dat wil ik niet zeggen".

#### Mediator: fysiek zwaar werk

De tweede mediator geeft de mate van fysiek zwaar werk van de respondent aan en heeft het label 'Mijn werk (is/was) fysiek veeleisend'. De variabele heeft de volgende antwoordcategorieën: 1 = "volledig mee oneens", 2 = "mee oneens", 3 = "mee eens", 4 = "volledig mee eens".

**Controlevariabelen: geslacht, leeftijd en werksector***Geslacht*

De eerste variabele waarvoor is gecontroleerd is geslacht. Hierbij geldt: 1 = "man", 2 = "vrouw". Van de variabele is een dummyvariabele gemaakt (zie Bijlage 1), waarbij geldt: 0 = "man", 1 = "vrouw".

*Leeftijd*

De tweede variabele waarvoor is gecontroleerd is leeftijd. De respondent kan zijn / haar leeftijd invullen. Van de variabele leeftijd zijn leeftijdscohorten gemaakt (zie Bijlage 1). De leeftijdscategorie 40-45 is leeftijdscohort 1 geworden, de leeftijdscategorie 46-50 leeftijdscohort 2, de leeftijdscategorie 51-55 leeftijdscohort 3, de leeftijdscategorie 56-60 leeftijdscohort 4, de leeftijdscategorie 61-65 leeftijdscohort 5 en tot slot de leeftijdscategorie 66-69 leeftijdscohort 6.

*Werksector*

De derde variabele waarvoor is gecontroleerd is de werksector van de respondent. De variabele heeft de volgende antwoordcategorieën: 1 = "landbouw, jacht, bosbouw, visserij", 2 = "winning van delfstoffen", 3 = "industrie", 4 = "productie en distributie van en handel in elektriciteit, aardgas, stoom en water", 5 = "bouwnijverheid", 6 = "handel (incl. reparatie van consumentenartikelen)", 7 = "horeca", 8 = "vervoer, opslag en communicatie", 9 = "financiële instellingen", 10 = "zakelijke dienstverlening (incl. onroerend goed, verhuur van roerende goederen)", 11 = "overheidsdiensten, openbaar bestuur en verplichte sociale verzekeringen", 12 = "onderwijs", 13 = "gezondheids- en welzijnszorg", 14 = "milieudienstverlening, cultuur, recreatie en overige dienstverlening", 15 = "overig".

Van de variabele zijn dummy's gemaakt (zie Bijlage 1). Van de antwoordcategorieën 1, 2, 3 en 5 is de dummy gemaakt die aangeeft of de respondent in de sector "industrie en bouw" werkt. Van de antwoordcategorieën 4, 6, 7, 8, 9 en 10 is een dummy gemaakt die aangeeft of de respondent in de sector "diensten en handel" werkt. Van de antwoordcategorieën 11, 12, 13 en 14 is een dummy gemaakt die aangeeft of de respondent in de "(semi)publieke sector" werkt. Antwoordcategorie 15 (overig) is de referentiecategorie.

### 3.4 Analyse-opzet

De hypothesen gaan over het verschil in opleidingsniveau en de mate van boosheid betreffende de stijging van de AOW-leeftijd. Deze hypothesen zijn getoetst aan de hand van een lineaire regressieanalyse. Er zijn verschillende modellen opgesteld om de hypothesen te kunnen toetsen. Eerst volgt het model met de afhankelijke variabele (boosheid) en de controlevariabelen geslacht, leeftijd en werksector. Dit is model 1.

Vervolgens is er gekeken in hoeverre er een verschil is tussen de opleidingsniveaus en de mate van boosheid. Ook zijn de controlevariabelen toegevoegd. Dit is model 2.

In model 3a en 3b zijn andere afhankelijke variabelen gebruikt. In model 3a is getoetst of opleidingsniveau significant verschilt in het inkomen. Ook zijn de controlevariabelen toegevoegd. In model 3b is getoetst of opleidingsniveau significant verschilt in de mate van fysiek zwaar werk. Ook zijn de controlevariabelen toegevoegd.

Aan model 2 is inkomen toegevoegd, waardoor er een nieuw model ontstaat, model 4a. Met behulp van dit model kan er worden onderzocht of het inkomen een deel verklaart van de mate waarin iemand boos is t.a.v. de stijging van de AOW-leeftijd.

Aan model 2 is vervolgens de mate van fysiek zwaar werk toegevoegd, waardoor er een nieuw model ontstaat, model 4b. Met behulp van dit model kan worden onderzocht of de mate van fysiek zwaar werk een deel verklaart van de mate waarin iemand boos is t.a.v. de stijging van de AOW-leeftijd.

Ten slotte is model 5 geschat, waar alle variabelen in zitten.

## 4. Resultaten

Dit hoofdstuk geeft de resultaten van het onderzoek weer. Allereerst volgen de beschrijvende statistieken van de variabelen. Vervolgens geeft het hoofdstuk de modevaluatie van de getoetste modellen weer en tot slot volgt de hypothesetoetsing van het onderzoek.

### 4.1 Beschrijvende statistieken

#### 4.1.1 Univariate verdelingen

Tabel 1: Univariate verdelingen van de variabelen (n=1259)

Variabele	Gemiddelde (standaarddeviatie) <sup>a</sup>	Minimum	Modus	Maximum
Boosheid	31,9% niet, 23,9% in geringe mate, 17,9% in redelijke mate, 26,3% in hoge mate	1,00	1,00	4,00
Opleidingsniveau	4,08 (1,33)	1,00	4,00	6,00
Inkomen	4,84 (1,95)	1,00	5,00	12,00
Fysiek werk	29,5% helemaal niet mee eens, 42,1% niet mee eens, 23,4% mee eens, 5,1% helemaal mee eens	1,00	2,00	4,00
Geslacht (man=0; vrouw=1)	51,7% man, 48,3% vrouw	0,00		1,00
Leeftijd	53,00	40,00	56,00	69,00
Werksector	9,78 (3,94)	1,00	13,00	15,00

<sup>a</sup>Bij nominale variabelen is de frequentieverdeling vermeld in percentages.

In tabel 1 zijn de univariate verdelingen van de variabelen weergegeven. Er is te zien dat de meeste respondenten niet boos zijn over de stijging van de AOW-leeftijd (31,9%). Gevolgd door de respondenten die in hoge mate boos zijn (26,3%).

Er is ook te zien dat het gemiddelde van opleidingsniveau 4,08 ( $SD=1,33$ ) is. Opleidingsniveau is gemeten in 6 categorieën en loopt van laag (basisonderwijs) naar hoog (WO). Het gemiddelde opleidingsniveau ligt rond het MBO.

Bij de variabele inkomen is het gemiddelde 4,84 ( $SD=1,95$ ). Inkomen loopt van een schaal van 1-12 waarbij 1 staat voor een inkomen van 500 euro of minder en 12 voor een inkomen hoger dan 7500 euro. Het gemiddelde netto-inkomen ligt tussen antwoordcategorie 4 (1501 t/m 2000 euro) en antwoordcategorie 5 (2001 t/m 2500 euro) in.

Bij de variabele fysiek zwaar werk is te zien dat de meeste respondenten aangeven dat ze hun werk fysiek niet veeleisend vinden (42,1%).

In tabel 1 staan ook de drie controlevariabelen weergegeven. Er is te zien dat er ongeveer evenveel mannen en vrouwen in de dataset zitten en dat de gemiddelde leeftijd van de respondenten 53 jaar is. Tot slot is de werksector van de respondent weergegeven, hieruit blijkt dat de modus van werksector antwoordcategorie 13 (de gezondheids- en welzijnszorg) is.

#### 4.1.2 Bivariate verdelingen

Tabel 2: Bivariate verdelingen van de variabelen ( $n=1259$ )

Variabelen	2.	3.	4.	5.	6.
1. Boosheid	-0,298** <sup>b</sup>	-0,224** <sup>b</sup>	0,146** <sup>c</sup>	0,116** <sup>c</sup>	0,202* <sup>b</sup>
2. Opleidingsniveau	-	0,379** <sup>a</sup>	-0,286** <sup>b</sup>	-0,019 <sup>a</sup>	-0,045 <sup>a</sup>
3. Inkomen		-	-0,311** <sup>b</sup>	-0,378** <sup>a</sup>	-0,013 <sup>a</sup>
4. Fysiek werk			-	0,121** <sup>c</sup>	0,126 <sup>b</sup>
5. Geslacht (0=man; 1=vrouw)				-	-0,030 <sup>a</sup>
6. Leeftijd					-

\*significant bij  $p < 0,005$ ; \*\*significant bij  $p < 0,001$ ; <sup>a</sup>= Pearson's correlation; <sup>b</sup>=VR Squared uit ANOVA; <sup>c</sup> = Cramer's V

In tabel 2 zijn de bivariate verdelingen van de variabelen te zien (voor een uitgebreidere uitleg per samenhang; zie Bijlage 2). De samenhang is berekend door een Pearson Correlation, een VR Squared uit ANOVA voor het verschil in gemiddelden en een Cramer's V. De samenhang tussen de variabele 'Werksector' en de andere variabelen staat toegelicht in bijlage 2 en is niet opgenomen in bovenstaande tabel.

In tabel 2 is te zien dat de mate van boosheid significant negatief samenhangt met opleidingsniveau ( $r = -0,298$ ;  $p < 0,001$ ). Een hogere opleiding gaat dus samen met minder boosheid ten aanzien van de stijging van de AOW-leeftijd. De samenhang is niet erg groot. Verder is te zien dat boosheid en inkomen significant negatief met elkaar samenhangen ( $r = -0,224$ ;  $p < 0,001$ ). Een hoger inkomen gaat samen met minder boosheid. Ook is te zien dat de mate van boosheid en de mate van fysiek zwaar werk significant positief met elkaar samenhangen ( $r = 0,146$ ;  $p < 0,001$ ). Het werk als meer fysiek zwaar ervaren gaat samen met het meer boos zijn omtrent de stijging van de AOW-leeftijd. De samenhang is niet erg groot.

Verder is te zien dat het inkomen en opleidingsniveau significant positief met elkaar samenhangen ( $r = 0,379$ ;  $p < 0,001$ ). Een hoger opleidingsniveau gaat dus gepaard met een hoger inkomen. Een hoger inkomen gaat ook significant gepaard met het werk minder fysiek zwaar vinden ( $r = -0,311$ ;  $p < 0,001$ ). Hetzelfde geldt voor de samenhang tussen het opleidingsniveau en de mate van

fysiek zwaar werk. Een hoger opleidingsniveau hangt significant negatief samen met in hoeverre het werk als fysiek zwaar wordt ervaren ( $r = -0,286$ ;  $p < 0,001$ ).

Ook valt uit de tabel te halen dat de mate van boosheid en geslacht zwak met elkaar samenhangen. De Cramer's V heeft een waarde van 0,116 (zie Bijlage 2). Tot slot hangen boosheid en leeftijd significant positief met elkaar samen ( $r = 0,202$ ;  $p < 0,005$ ). Dus een hogere leeftijd hangt significant samen met het meer boos zijn betreffende de stijging van de AOW-leeftijd.

## 4.2 Modevaluatie

### Kwaliteit van het model

De kwaliteit van het model is te beoordelen aan de hand van de  $R^2_{\text{change}}$ . De  $R^2_{\text{change}}$  geeft de verandering van  $R^2$  aan wanneer er een nieuwe voorspeller bij komt. In tabel 3 van paragraaf 4.3 is te zien dat de  $R^2_{\text{change}}$  steeds een beetje stijgt naarmate er meer variabelen aan het model worden toegevoegd. Verder is te zien in tabel 3 dat model 5 een verklaarde variantie heeft van 10,4%. De verklaarde variantie is hier hoger dan in model 2 (6,8%), met alleen de controlevariabelen en de onafhankelijke variabele als voorspeller van de afhankelijke variabele. Het volledige model verklaart significant de meeste variantie. Een verklaarde variantie van 10,4% is redelijk laag en er is dus nog een groot deel onverklaarde variantie.

### Multicollineariteit

Multicollineariteit houdt in dat voorspellers te sterk met elkaar samenhangen. Een hulpmiddel om de multicollineariteit te meten is de VIF. Over het algemeen geldt dat wanneer de VIF lager is dan 4, er geen sprake is van ernstige multicollineariteit. In tabel 4 is in de laatste rij de VIF van elke variabele weergegeven. Alle variabelen hebben een VIF lager dan 4. Hierdoor kan er worden geconcludeerd dat er geen sprake lijkt te zijn van multicollineariteit.

### Modelassumpties

Om te kijken of het model geschikt is voor de data die is gebruikt en of de resultaten kunnen worden gegeneraliseerd naar de populatie, wordt er gekeken of er aan de volgende vier modelassumpties wordt voldaan.

De eerste assumptie stelt dat de observaties onafhankelijk van elkaar moeten zijn. Dit betekent dat de respondenten onafhankelijk van elkaar zijn getrokken in de steekproef, waardoor er geen samenhang zal ontstaan tussen de scores van de respondenten. De respondenten zijn aselekt getrokken op basis van een steekproef van het Centraal Bureau van de Statistiek. Echter, het LISS panel maakt gebruik van geneste data, het is gebaseerd op een steekproef van huishoudens. De waarnemingen van individuele personen in het LISS panel zijn geclusterd in huishoudens, zo kunnen



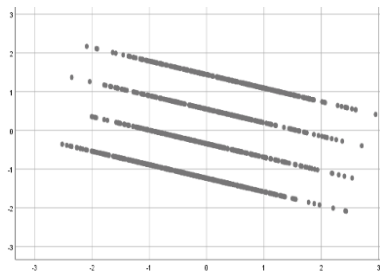
er meerdere individuen van hetzelfde huishouden meedoen aan het panel. Hierdoor zijn de waarnemingen niet volledig onafhankelijk van elkaar. Echter, het is een groot panel wat inhoudt dat de personen uit hetzelfde huishouden niet perse dezelfde vragenlijsten invullen (LISS, z.d.). Hierdoor zijn de waarnemingen ook niet volledig afhankelijk van elkaar. Er is sprake van een gedeeltelijke schending van deze assumptie.

De tweede assumptie stelt dat er een lineair verband moet zijn tussen de afhankelijke variabele en de onafhankelijke variabelen. Om te kijken of dit het geval is, wordt er gekeken naar de spreidingsdiagram weergegeven in figuur 2. Op de x-as staat de variabele die de mate van boosheid van de respondent aangeeft. Het gemiddelde voor elke set van x-waarden moet ongeveer gelijk aan 0 zijn. Dit betekent dat de punten ongeveer gelijk rondom de nullijn liggen. Dit is niet het geval. De punten in de spreidingsdiagram boven de lijn en de punten onder de lijn heffen elkaar niet op en het gemiddelde voor elke set van x-waarden is niet gelijk aan nul. Er zijn duidelijk vier categorieën in te zien. Er kan dus niet worden gesteld dat er een lineair verband is tussen de residuen. Deze assumptie kan niet worden aangenomen.

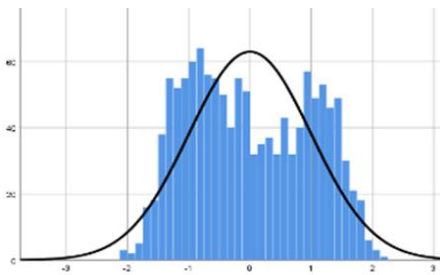
De derde assumptie stelt dat er sprake moet zijn van homoscedasticiteit (een constante conditionele standaarddeviatie). Dit houdt in dat voor elke set van waarden van de x-en, y dezelfde conditionele standaarddeviatie heeft (in de populatie). Dit is het geval wanneer de residuen (op de x-as) een gelijke variantie hebben. De gelijke variantie houdt in dat de residuen gemiddeld dezelfde afstand hebben tot de regressielijn. Om de homoscedasticiteit te bepalen wordt er wederom gekeken naar de spreidingsdiagram in figuur 2. De spreiding is overal ongeveer even groot. De residuen liggen in een redelijk gelijke band rondom de nullijn.

De laatste assumptie stelt dat de conditionele verdeling van y normaal verdeeld moet zijn. Dit wordt gecontroleerd door middel van het histogram van de gestandaardiseerde residuen en de PP-plot, weergegeven in figuur 3 en 4. In de histogram is te zien dat staafjes niet de normaalkromme volgen. De verdeling is niet normaal verdeeld. Verder is in de PP-plot te zien dat de punten niet de zwarte lijn volgen. Wanneer de verdeling normaal verdeeld is zullen de punten in de figuur gelijk zijn aan de zwarte lijn, dit is niet het geval. Er is een lichte s-curve te zien. Dit betekent dat er pieken in de data voorkomen, wat terug te zien is in de histogram. Er kan dus gesteld worden dat er niet aan de normale verdeling wordt voldaan. Hierdoor wordt de assumptie geschonden. Dit schaadt vooral de power, waardoor de gevonden p-waarden onbetrouwbarder worden en er een kans bestaat dat de hypothesen onterecht worden aangenomen.

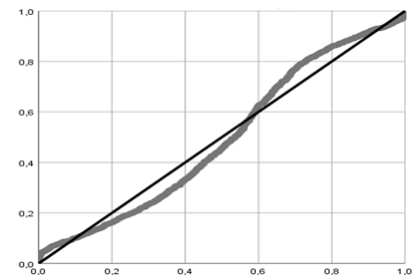
In totaal zijn er drie (waarvan één assumptie gedeeltelijk) van de vier modelassumpties geschonden, onder andere door het gebruik van geneste data en door een niet-normaal verdeelde afhankelijke variabele. Hierdoor moet er rekening worden gehouden met de interpretatie en generalisatie van de data.



Figuur 2: Residual plot



Figuur 3: Histogram met normaalkromme



Figuur 4: PP-plot

## Uitbijters

Het is ten slotte nog van belang om te kijken of er uitbijters in de data voorkomen die een eventueel vertekend beeld kunnen geven. Er wordt door middel van de Leverage, de Cook's distance, DfBeta en DfFit gekeken welke punten de uitbijters of invloedrijke punten zijn. Er wordt specifiek gekeken naar de respondenten die op meerdere maten opvallen.

Eerst wordt er gekeken naar invloedrijke punten in de x-richting: de Leverage. Met behulp van de formule  $3 * P/N$  is te controleren of er een probleem kan ontstaan. Hierbij is  $P(8)$  het aantal geschatte parameters en  $N(1259)$  het totale aantal respondenten. Dit geeft een Leverage-grens van: 0,01906. In de dataset wordt gekeken naar waarden die boven deze grens liggen.

De Cook's distance is een combinatie van outlier in de y-richting en de x-richting. De vuistregel hierbij is dat wanneer de Cook's distance groter is dan 1, dit een probleem is. Dit is niet het geval. Er wordt een grenswaarde uitgerekend ( $4/N = 0,0032$ ). Er wordt gekeken naar waarden die boven deze grens liggen.

Met de DfBeta kijk men hoe ver de coëfficiënt verandert als de regressieanalyse opnieuw wordt gedaan zonder de betreffende case. De DfBeta wordt voor elke parameter uitgerekend. Er is gekeken naar de grootste positieve en negatieve DfBeta's die voor de grootste invloed kunnen zorgen.

Met de DfFit wordt gekeken hoe ver y verandert als de regressieanalyse opnieuw wordt uitgevoerd zonder de betreffende case. De grootste negatieve DfFit waarde is -0,02950 en de grootste positieve DfFit waarde is 0,04891. Het zijn erg kleine waarden, wanneer de betreffende cases worden verwijderd uit de dataset zal de helling dan ook niet erg worden beïnvloed.

In bijlage 3 staat een uitgebreidere toelichting van de uitbijters. De betreffende cases die op meerdere maten hoog scoren zijn uit de dataset gehaald. Er zijn geen opmerkelijke veranderingen te zien in de hellingen of de significantie van de variabelen (zie Bijlage 3).

### 4.3 Hypothesetoetsing

Tabel 3: Parameters uit de regressieanalyse met de 9 variabelen in 5 modellen met als afhankelijke variabele de mate van boosheid

	Model 1		Model 2		Model 4a		Model 4b		Model 5		VIF
	B (SE)	p	B (SE)	P	B (SE)	p	B (SE)	p	B (SE)	p	
Constante	2,435 (0,132)	<0,001	3,359 (0,160)	<0,001	3,610 (0,170)	<0,001	2,682 (0,190)	<0,001	2,931 (0,223)	<0,001	
Geslacht (0=man)	0,056 (0,071)	0,436	0,006 (0,069)	0,934	-0,107 (0,074)	0,147	-0,010 (0,068)	0,883	-0,092 (0,073)	0,206	1,320
Leeftijd	-0,026 (0,023)	0,259	-0,040 (0,023)	0,075	-0,041 (0,022)	0,069	-0,035 (0,022)	0,114	-0,036 (0,022)	0,103	1,028
Werksector											
D1-industriebouw	0,156 (0,133)	0,239	0,205 (0,128)	0,110	0,249 (0,128)	0,052	0,189 (0,126)	0,136	0,223 (0,126)	0,078	2,274
D2-diensthandel	-0,058 (0,121)	0,630	0,022 (0,117)	0,848	0,039 (0,116)	0,738	0,058 (0,115)	0,615	0,067 (0,115)	0,560	2,715
D3-semipubliek	-0,010 (0,116)	0,934	0,175 (0,113)	0,124	0,188 (0,113)	0,096	0,155 (0,112)	0,165	0,167 (0,111)	0,135	3,060
Opleidingsniveau			-0,237 (0,025)	<0,001	-0,192 (0,027)	<0,001	-0,197 (0,025)	<0,001	-0,167 (0,027)	<0,001	1,276
Inkomen					-0,082 (0,019)	<0,001			-0,061 (0,020)	0,002	1,454
Fysiek werk							0,248 (0,039)	<0,001	0,225 (0,039)	<0,001	1,121
$R^2$ (adjusted)	0,001		0,068		0,081		0,097		0,104		
$R^2$ change	0,005	0,301	0,068	<0,001	0,013	<0,001	0,030	<0,001	0,023 <sup>a</sup> 0,007 <sup>b</sup>	<0,001 0,002	
n	1259		1259		1259		1259		1259		

<sup>a</sup>  $R^2$  change ten opzichte van model 4a; <sup>b</sup>  $R^2$  change ten opzichte van model 4b

Tabel 4: Parameters uit de regressieanalyse met als afhankelijke variabele inkomen (3a) en de mate van fysiek zwaar werk (3b)

	Model 3a <sup>a</sup>		Model 3b <sup>b</sup>	
	B (SE)	p	B (SE)	P
Constante	3,066 (0,231)	<0,001	2,732 (0,115)	<0,001
Geslacht (0=man)	-1,371 (0,100)	<0,001	0,064 (0,050)	0,202
Leeftijd	-0,008 (0,033)	0,811	-0,020 (0,016)	0,213
Werksector				
D1-industriebouw	0,534 (0,185)	0,004	0,068 (0,093)	0,463
D2-diensthandel	0,202 (0,169)	0,232	-0,144 (0,084)	0,088
D3-semipubliek	0,160 (0,164)	<0,329	0,079 (0,082)	0,334
Opleidingsniveau	0,549 (0,036)	<0,001	-0,162 (0,018)	<0,001
$R^2$ (adjusted)	0,284		0,071	
n	1259		1259	

<sup>a</sup> afhankelijke variabele is inkomen; <sup>b</sup> afhankelijke variabele is fysiek zwaar werk

In tabel 3 is de regressieanalyse weergegeven met als afhankelijke variabele de mate van boosheid ten aanzien van de stijging van de AOW-leeftijd. In model 1 zijn de controlevariabelen toegevoegd aan het model met de afhankelijke variabele. Hier is te zien dat vrouwen bozer zijn over de stijging van de AOW-leeftijd dan mannen ( $b= 0,056$ ;  $p=0,436$ ), gecontroleerd voor de overige variabelen. Ook blijkt dat men bozer is naarmate men in een lager leeftijdscohort zit ( $b= -0,026$ ;  $p=0,259$ ), gecontroleerd voor de overige variabelen. Wanneer men in de industrie en bouw werkt, ten opzichte van de overige categorie (de referentiecategorie), zal de boosheid van de respondent met 0,156 toenemen. Wanneer men in de diensten en handel werkt zal de boosheid ten opzichte van de overige categorie, afnemen met 0,058. Wanneer men in de semipublieke sector werkt zal de boosheid ten opzichte van de overige categorie afnemen met 0,010. De hellingen van de dummy's zijn redelijk klein, en dus is het effect niet sterk. Er lijkt geen verschil in de invloed van de werksectoren te zijn, omdat de dummy's niet statistisch significant zijn.

Door middel van model 2 kan er worden gekeken of de eerste hypothese wordt ondersteund door de resultaten. De eerste hypothese luidt als volgt: *Hoe hoger het opleidingsniveau van de respondent, hoe minder boos men is over de stijging van de AOW-leeftijd.* In model 4 is te zien dat er een statistisch significant verschil is in de mate van boosheid voor de verschillende opleidingsniveaus ( $b= -0,237$ ;  $p<0,001$ ), gecontroleerd voor de overige variabelen. Dus hoe hoger het opleidingsniveau van de respondent, hoe minder boos men is over de stijging van de AOW-leeftijd. De mate van boosheid van de respondent is gemeten op een schaal van 1-4, het effect van opleidingsniveau op de mate van boosheid is dus niet heel groot. Wel is het effect statistisch significant.

In model 3a, weergegeven in tabel 4, is de afhankelijke variabele het inkomen. Hieruit blijkt dat er een significante helling is van opleidingsniveau ( $b= 0,549$ ;  $p<0,001$ ), gecontroleerd voor de overige variabelen. Dit betekent hoe hoger het opleidingsniveau van de respondent, hoe hoger het inkomen van de respondent. De helling is redelijk klein, het inkomen van de respondent is gemeten op een schaal van 1-12.

In model 3b, weergegeven in tabel 4, is de afhankelijke variabele de mate van fysiek zwaar werk. Er is te zien dat er een significante helling is van het opleidingsniveau ( $b= -0,162$ ;  $p<0,001$ ), gecontroleerd voor de overige variabelen. Hieruit blijkt hoe hoger het opleidingsniveau van de respondent, hoe minder fysiek zwaar werk de respondent heeft. De helling is redelijk klein, de mate van fysiek zwaar werk is gemeten op een schaal van 1-4.

Aan de hand van de modellen 4a, 4b en 5 is te onderzoeken of er sprake is van een partieel mediërend effect van het inkomen op de mate van boosheid en de mate van fysieke zwaar werk op de mate van boosheid. Aan de hand van model 4a kan er worden gekeken of de tweede hypothese wordt ondersteund door de resultaten. De tweede hypothese luidt als volgt: *Het negatieve effect van het hebben van een hoger opleidingsniveau op de mate van boosheid t.a.v. de stijging van de AOW-*

*leeftijd, wordt (gedeeltelijk) verklaard door inkomen.* Model 4a laat zien dat wanneer men een hoger inkomen heeft, men minder boos is over de stijging van de AOW-leeftijd ( $b = -0,082$ ;  $p < 0,001$ ), gecontroleerd voor de overige variabelen. Verder laat model 4a zien dat de helling van de onafhankelijke variabele opleidingsniveau is afgenomen in vergelijking met model 2 (van  $-0,237$  naar  $-0,192$ ), waardoor inkomen een deel verklaart van het effect van opleidingsniveau op de mate van boosheid. De afname is echter niet statistisch significant, omdat de betrouwbaarheidsintervallen van opleidingsniveau elkaar grotendeels overlappen in model 2 en model 4a (zie Bijlage 2).

Met model 4b kan er worden gekeken of de derde hypothese wordt ondersteund door de resultaten. De derde hypothese luidt als volgt: *Het negatieve effect van het hebben van een hoger opleidingsniveau op de mate van boosheid t.a.v. de stijging van de AOW-leeftijd, wordt (gedeeltelijk) verklaard door de mate van fysiek zwaar werk.* Model 4b laat zien dat wanneer men meer fysiek zwaar werk heeft, men bozer is over de stijging van de AOW-leeftijd ( $b = 0,248$ ;  $p < 0,001$ ), gecontroleerd voor de overige variabelen. Verder laat model 4b zien dat de helling van opleidingsniveau is afgenomen in vergelijking met model 2 (van  $-0,237$  naar  $-0,197$ ), waardoor de fysieke zwaarte van het werk een deel verklaart van het effect van opleidingsniveau op de mate van boosheid. De afname is echter niet statistisch significant, omdat de betrouwbaarheidsintervallen van opleidingsniveau in model 2 en model 4b elkaar grotendeels overlappen (zie Bijlage 2).

Door middel van model 5 kan er eveneens worden gekeken of er sprake is van een mediatie-effect, maar dan wanneer de beide mediators, inkomen en fysieke zwaarte van het werk, worden toegevoegd aan model 2. Model 5 laat zien dat de helling van opleidingsniveau kleiner wordt in vergelijking met model 2 (van  $-0,237$  naar  $-0,167$ ), wat inhoudt dat het inkomen van de respondent en de fysieke zwaarte van het werk een deel van de invloed (opleidingsniveau op boosheid) verklaren. De afname is echter niet statistisch significant, doordat de betrouwbaarheidsintervallen van opleidingsniveau elkaar deels overlappen in model 2 en model 5 (zie Bijlage 2).

## 5. Conclusie en discussie

In dit onderzoek is antwoord gegeven op de vraag wie er boos zijn betreffende de stijging van de AOW-leeftijd. Uit het onderzoek is gebleken dat mensen met een hoger opleidingsniveau gemiddeld gezien minder boos zijn over de stijging van de AOW-leeftijd dan mensen met een lager opleidingsniveau (in lijn met hypothese 1). De resultaten geven echter wel aan dat het effect klein is. De tweede hypothese luidt als volgt: *Het negatieve effect van het hebben van een hoger opleidingsniveau op de mate van boosheid t.a.v. de stijging van de AOW-leeftijd, wordt (gedeeltelijk) verklaard door inkomen.* Deze hypothese kan niet worden aangenomen, omdat de resultaten niet significant bevonden zijn. Tot slot volgt de derde hypothese: *Het negatieve effect van het hebben van een hoger opleidingsniveau op de mate van boosheid t.a.v. de stijging van de AOW-leeftijd, wordt (gedeeltelijk) verklaard door de mate van fysiek zwaar werk.* De derde hypothese kan eveneens niet worden aangenomen, omdat de resultaten niet significant bevonden zijn.

Er is gebleken dat er boosheid heerst rondom de stijging van de AOW-leeftijd. Het doel van dit onderzoek was om deze boosheid te verklaren, enerzijds door te kijken wat de invloed is van opleidingsniveau op de mate van boosheid t.a.v. de stijging van de AOW-leeftijd en anderzijds om meer diepte te geven aan de invloed van opleidingsniveau op de mate van boosheid. Deze diepte is aan het onderzoek gegeven door ook naar het inkomen en de fysieke zwaarte van het werk te kijken. Er kan niet worden geconcludeerd dat het inkomen en de fysieke zwaarte van het werk een rol spelen in de invloed van opleidingsniveau op de mate van boosheid t.a.v. de stijging van de AOW-leeftijd.

Het is gebleken dat vooral laag- en middelbaar opgeleide werknemers frustratie ervaren ten aanzien van de stijging van de AOW-leeftijd. Dit is in tegenstelling tot universitair geschoolde werknemers, die minder negatieve emoties ervaren ten aanzien van de stijging (De Beer, Van Dalen & Henkens, 2017a). Dit is ook teruggevonden in deze studie. Verder is gebleken dat mensen met een hoger inkomen minder boosheid kunnen ervaren rondom de stijging van de AOW-leeftijd, onder andere doordat ze meer financiële vrijheid hebben (Kok et al., 2017). In dit onderzoek is uit de resultaten gekomen dat naarmate het inkomen hoger is men minder boos is omtrent de stijging van de AOW-leeftijd, maar inkomen medieert niet de invloed van opleidingsniveau op boosheid. Verder is gebleken dat lageropgeleiden meer fysiek zwaar werk hebben (Kok et al., 2017) en een verminderde *employability* ervaren in het werk (Smit, Andriessen & Stark, 2005), waardoor de frustratie ook groter kan zijn rondom de stijging van de AOW-leeftijd. In dit onderzoek is gevonden dat lageropgeleiden meer fysiek zwaar werk uitoefenen. Er kan echter niet worden gezegd dat fysiek zwaar werk een deel van de invloed van opleiding op de mate van boosheid verklaart, omdat het niet significant is bevonden.

De tweede en derde hypothese kunnen niet worden ondersteund door de resultaten. Dit zou kunnen liggen aan een aantal beperkingen van het onderzoek. Allereerst zou de schending van de modelassumpties een verklaring kunnen zijn. Drie (waarvan één gedeeltelijk) van de vier modelassumpties zijn namelijk geschonden. De eerste assumptie is gedeeltelijk geschonden doordat de gebruikte dataset gebruikt maakt van geneste data. De waarnemingen van individuele personen in het LISS panel zijn geclusterd in huishoudens, waardoor er sprake van afhankelijkheid kan voorkomen in de dataset. Voor vervolgonderzoek is het daarom belangrijk om, wanneer er sprake is van geneste data, een multilevel analyse uit te voeren. Hierdoor kan de onafhankelijkheid van het onderzoek worden gewaarborgd. Ook zijn de tweede en vierde assumptie geschonden, wat onder andere komt doordat de afhankelijke variabele niet normaal verdeeld is. Hierdoor is een lineaire regressieanalyse eigenlijk ongeschikt. Toch is besloten om hiermee door te gaan en om strenger te toetsen, zodat het model alsnog kan worden gebruikt. Er is een significantieniveau van 0,001 gebruikt in plaats van het meer gebruikelijke significantieniveau van 0,005.

De vraagstelling van de afhankelijke variabele kan ook een oorzaak zijn voor het vinden van de (niet-significante) resultaten. Zo is de boosheid van de respondent gemeten aan de hand van een enkele vraag die luidt: *'Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lang moet doorwerken?'*. Alhoewel de meeste mensen wel weten wat boos zijn inhoudt, kan men de emotie 'boosheid' anders interpreteren. Het is wellicht beter om aan de hand van meerdere vragen er achter te komen hoe de respondent zich voelt omtrent het langer moeten doorwerken. Zo kunnen er vragen worden gesteld die vragen naar in hoeverre de respondent het fijn vindt om langer door te moeten werken. Ook kan er nog worden gevraagd naar bijvoorbeeld verdriet, angst en ergernis rondom de stijging van de AOW-leeftijd. Hierdoor kan er wellicht een beter beeld van de negatieve emoties van de werkende Nederlandse bevolking, omtrent de stijging van de AOW-leeftijd, in beeld worden gebracht.

Een andere beperking in dit onderzoek is dat de populatie uit een klein deel van de werkende bevolking bestaat, namelijk de leeftijdscategorie 40-69 jaar. Wanneer er naar de gehele werkende bevolking wordt gekeken, zouden er misschien wel andere (significante) resultaten naar voren komen. De AOW-leeftijd voor jongere leeftijdscategorieën ligt hoger, hierdoor zouden zij wellicht meer boosheid kunnen ervaren rondom de stijging van de AOW-leeftijd. Echter zou het ook kunnen dat jongere mensen minder bezig zijn met de stijging van de AOW-leeftijd, waardoor ze weinig of geen boosheid ervaren. Voor vervolgonderzoek is het daarom interessant om naar een grotere populatie te kijken, waar ook individuen jonger dan 40 jaar in zijn meegenomen.

Er is minimaal bewijs gevonden in deze studie dat mensen met een hoger opleidingsniveau minder boos zijn over de stijging van de AOW-leeftijd dan mensen met een lager opleidingsniveau. Er

is echter geen ondersteuning gevonden dat de mate van fysiek zwaar werk en het inkomen een deel van deze invloed verklaren. Al met al is het belangrijk dat er dieper op de invloed van opleidingsniveau op de mate van boosheid wordt ingegaan in verder onderzoek. Ook kan er verder worden gekeken naar het inkomen en de mate van fysiek zwaar werk, omdat uit de literatuur wel is gebleken dat deze twee factoren een rol kunnen spelen. In het onderzoek is sprake van een groot deel onverklaarde variantie. Dit is niet vreemd, want er zullen altijd andere oorzaken buiten het model zitten die de mate van boosheid verklaren. Zo kan er in vervolgonderzoek ook naar andere relevante factoren worden gekeken (bijvoorbeeld; het hebben van een stressvolle baan) die de invloed van opleidingsniveau op de mate van boosheid kunnen verklaren, zodat er hopelijk beter in beeld wordt gebracht wie er precies boos zijn zodat hier iets aan kan worden gedaan.



## Literatuur

- Bruggink, J. W. (2009). Ontwikkelingen in (gezonde) levensverwachting naar opleidingsniveau. *Bevolkingstrends*, 71-75.
- Bruggink, J.W., M.J. Garssen, B.J.H. Lodder & M. Kardal (2009). Trends in gezonde levensverwachting. *Bevolkingstrends*, 57(1), 60–66.
- Damman, M., & Oude Mulders, J. (2019). Doorwerken tot op hogere leeftijd, wat vinden werkgevers, werknemers en zzp'ers? *Network for Studies on Pensions, Aging and Retirement (NETSPAR)*, 1-22.
- De Beer, J. A. A., Van Dalen, H. P., & Henkens, K. (2017a). Wanneer voelt de verhoging van de AOW leeftijd als diefstal? *Me Judice*, 9, 1-4.
- De Beer, J. A. A., & Van der Gaag, N. (2019). Stijging van de AOW-leeftijd: Nieuw akkoord, einde discussie? *Tijdschrift voor Arbeidsvraagstukken*, 35(4), 370-381.
- De Beer, J. A. A., Van Dalen, H. P., & Henkens, K. (2017b). Stijgt de AOW-leeftijd niet te hard? *Me Judice*. <http://www.mejudice.nl/artikelen/detail/stijgt-de-aowleeftijd-niet-te-hard>
- Henkens, K., Van Solinge, H., Damman, M., & Dingemans, E. (2016). Langer doorwerken valt nog niet mee. *Demos Bulletin over Bevolking en Samenleving*, 32(2), 1-4.
- Hoeymans, N. (2011). Kan iedereen langer doorwerken? *Tijdschrift voor Gezondheidswetenschappen*, 2, 71-72. <https://doi.org/10.1007/s12508-011-0024-2>
- Kok, L., Kroon, L., Lammers, M., van Soest, A., & Weel ter, B. (2017). *Gevolgen Flexibele AOW-Leeftijd*. SEO Economisch Onderzoek. Geraadpleegd op 30 maart 2022, van

[https://pure.uvt.nl/ws/portalfiles/portal/31783609/2017\\_48\\_Gevolgen\\_flexibele\\_AOW\\_leeftijd\\_02.pdf](https://pure.uvt.nl/ws/portalfiles/portal/31783609/2017_48_Gevolgen_flexibele_AOW_leeftijd_02.pdf)

LISS PANEL. (z.d.). – [website.lisspanel.nl](http://website.lisspanel.nl)

Mol, M., & De Vries, J. (2009). *Ziekteverzuim het laagst bij werknemers met een hoge mate van autonomie en veel steun van collega's en leidinggevenden* (CBS). Geraadpleegd van <https://www.cbs.nl/nl-nl/achtergrond/2009/22/ziekteverzuim-het-laagst-bij-werknemers-met-een-hoge-mate-van-autonomie-en-veel-steun-van-collega-s-en-leidinggevenden>

Mulder, L. (2018). Lente-artikel szw: lang zullen we leven ... en wat betekent dit voor werk? *Tijdschrift Voor Arbeidsvraagstukken*, 34(1). <https://doi.org/10.5117/2018.034.001.002>

Smulders, P., Houtman, I., & Van Den Bossche, S. (2009). *ESB Economisch Statistische Berichten*, 94, 682-684

Stam, S., M.J. Garssen, M. Kardal & B.J.H. Lodder (2008). *Hoogopgeleiden leven lang en gezond* (CBS). Geraadpleegd van <https://www.cbs.nl/-/media/imported/documents/2008/44/2008-c156-pub.pdf?la=nl-nl>

Visser, M. (2019). Ervaringen met en opvattingen over langer doorwerken. *Tijdschrift Voor Arbeidsvraagstukken*, 35, 382–394.

## Bijlage 1 – Overzicht variabelen

### Filteren afhankelijke variabele 'Boosheid'

De mate van boosheid van de respondent is bevraagd aan een deel van de respondenten. Er is gefilterd op de afhankelijke variabele die de boosheid ten aanzien van de stijging van de AOW-leeftijd aangeeft.

*Syntax:* \*filteren op de mate van boosheid.

```
*filterboosheid.
```

```
COMPUTE filterv3d=v3d.
```

```
EXECUTE.
```

```
FILTER OFF.
```

```
USE ALL.
```

```
SELECT IF (NOT(filterv3d=0)).
```

```
EXECUTE.
```

**Variabele 'Boosheid'**

**1.1 Oorspronkelijke variabele**

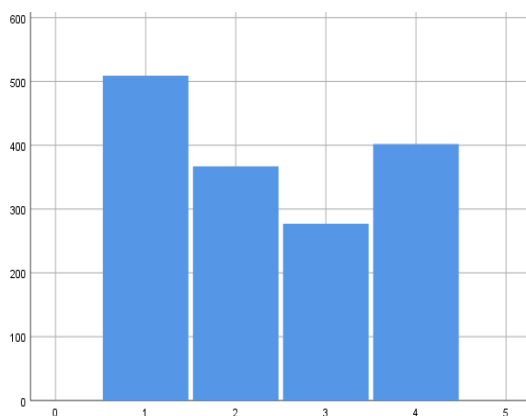
De afhankelijke variabele geeft de boosheid ten aanzien van de stijging van de AOW-leeftijd weer. Hierbij wordt gevraagd: 'Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lang moet doorwerken?' De variabele heeft de volgende antwoordcategorieën: 1 = "niet", 2 = "in geringe mate", 3 = "in redelijke mate", 4 = "in hoge mate". De meeste respondenten geven aan niet boos te zijn over de stijging van de AOW-leeftijd, dit is te zien in de frequentietabel hieronder. Er zijn 1555 respondenten die de vraag hebben beantwoord en er zijn geen missings. In de figuur hieronder is de verdeling van de mate van boosheid te zien. De verdeling is niet normaal verdeeld.

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness	Kurtosis
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic
boosheid Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan	1555	1	4	2,37	1,185	1,405	,192	,062
Valid N (listwise)	1555							

**boosheid Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Niet	509	32,7	32,7	32,7
2 In geringe mate	367	23,6	23,6	56,3
3 In redelijke mate	277	17,8	17,8	74,1
4 In hoge mate	402	25,9	25,9	100,0
Total	1555	100,0	100,0	



**Syntax: \*beschrijvende statistieken boosheid.**

```
DESCRIPTIVES VARIABLES=v3d
  /STATISTICS=MEAN STDDEV VARIANCE MIN MAX KURTOSIS SKEWNESS.
```

**Syntax: \*frequenties boosheid.**

```
FREQUENCIES VARIABLES=v3d
  /NTILES=4
```

```
/NTILES=10  
/STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT  
/HISTOGRAM  
/ORDER=ANALYSIS.
```

**Variabele 'Opleidingsniveau'****2.1 Oorspronkelijke variabele**

De onafhankelijke variabele geeft het opleidingsniveau van de respondent aan. Bij deze variabele is het label 'Opleiding in CBS-categorieën'. De variabele heeft de volgende antwoordcategorieën: 1 = "basisonderwijs", 2 = "vmbo", 3 = "havo/vwo", 4 = "mbo", 5 = "hbo", 6 = "wo", 9 = "Onbekend (missing)". De verdeling van de variabele is redelijk normaal. Er is een oververtegenwoordiging van respondenten van het mbo en het hbo. Dit is te zien in de frequentietabel. Er zijn in totaal 1552 respondenten die de vraag hebben beantwoord en er zijn 3 missings.

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
oplcac Level of education in CBS (Statistics Netherlands) categories	1552	1	6	4,07	1,324	-,472	,062	-,619	,124
Valid N (listwise)	1552								

**oplcac Level of education in CBS (Statistics Netherlands) categories**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 basisonderwijs	40	2,6	2,6	2,6
	2 vmbo	242	15,6	15,6	18,2
	3 havo/vwo	140	9,0	9,0	27,2
	4 mbo	472	30,4	30,4	57,6
	5 hbo	458	29,5	29,5	87,1
	6 wo	200	12,9	12,9	100,0
	Total	1552	99,8	100,0	
Missing	System	3	,2		
Total		1555	100,0		

**Syntax: \*beschrijvende statistieken opleidingsniveau.**

```
DESCRIPTIVES VARIABLES=oplcac
  /STATISTICS=MEAN STDDEV VARIANCE MIN MAX KURTOSIS SKEWNESS.
```

**Syntax: \*frequencies opleidingsniveau.**

```
FREQUENCIES VARIABLES=oplcac.
  /NTILES=4
  /NTILES=10
  /STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT
  /HISTOGRAM
  /ORDER=ANALYSIS.
```

**2.1 Bewerking variabele**

De missings van de variabele worden verwijderd en ook zal antwoordcategorie 9 worden verwijderd.

**Syntax:** \*verwijderen missings opleidingsniveau.

```
FILTER OFF.
USE ALL.
SELECT IF (NOT(oplcat=0)).
EXECUTE.
```

**2.3 Uiteindelijke variabele**

De beschrijvende statistieken zien er na de bewerking hetzelfde uit.

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
oplcat Level of education in CBS (Statistics Netherlands) categories	1552	1	6	4,07	1,324	-,472	,062	-,619	,124
Valid N (listwise)	1552								

**Syntax:** \*beschrijvende statistieken opleidingsniveau.

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

**Variabele 'Inkomen'****3.1 Oorspronkelijke variabele**

De eerste mediator geeft het inkomen van de respondent aan met het label 'Persoonlijk netto maandinkomen'. De variabele heeft de volgende antwoordcategorieën: 0 = "geen inkomen", 1 = "verdient 500 euro of minder", 2 = "verdient 501 t/m 1000 euro", 3 = "verdient 1001 t/m 1500 euro", 4 = "verdient 1501 t/m 2000 euro", 5 = "verdient 2001 t/m 2500 euro", 6 = "verdient 2501 t/m 3000 euro", 7 = "verdient 3001 t/m 3500 euro", 8 = "verdient 3501 t/m 4000 euro", 9 = "verdient 4001 t/m 4500 euro", 10 = "verdient 4501 t/m 5000 euro", 11 = "verdient 5001 t/m 7500 euro", 12 = "Meer dan 7500 euro", 13 = "Dat weet ik echt niet", 14 = "Dat wil ik niet zeggen". Hieronder zijn de beschrijvende statistieken en de frequentietabel van de variabele te zien. De meeste respondenten geven aan tussen de 2001-2500 euro te verdienen. De variabele is niet normaal verdeeld, er is een piek bij antwoordcategorie 0 die ver boven de normale curve uitsteekt.

**nettocat Personal net monthly income in categories**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 geen inkomen	17	1,1	1,1	1,1
	1 EUR 500 of minder	32	2,1	2,1	3,2
	2 EUR 501 t/m EUR 1000	110	7,1	7,1	10,2
	3 EUR 1001 t/m EUR 1500	191	12,3	12,3	22,6
	4 EUR 1501 t/m EUR 2000	328	21,1	21,1	43,7
	5 EUR 2001 t/m EUR 2500	329	21,2	21,2	64,9
	6 EUR 2501 t/m EUR 3000	213	13,7	13,7	78,6
	7 EUR 3001 t/m EUR 3500	110	7,1	7,1	85,7
	8 EUR 3501 t/m EUR 4000	63	4,1	4,1	89,8
	9 EUR 4001 t/m EUR 4500	29	1,9	1,9	91,6
	10 EUR 4501 t/m EUR 5000	8	,5	,5	92,1
	11 EUR 5001 t/m EUR 7500	17	1,1	1,1	93,2
	12 Meer dan EUR 7500	7	,5	,5	93,7
	13 Dat weet ik echt niet	19	1,2	1,2	94,9
	14 Dat wil ik niet zeggen	79	5,1	5,1	100,0
Total	1552	99,8	100,0		
Missing	System	3	,2		
Total		1555	100,0		

**Syntax: \*beschrijvende statistieken inkomen.**

```
DESCRIPTIVES VARIABLES=nettocat
  /STATISTICS=MEAN STDDEV VARIANCE MIN MAX KURTOSIS SKEWNESS.
```

**Syntax: \*frequenties inkomen.**

```
FREQUENCIES VARIABLES=nettocat
  /NTILES=4
  /NTILES=10
  /STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT
  /HISTOGRAM
  /ORDER=ANALYSIS.
```



**3.2 Bewerking variabele**

Het aantal respondenten dat de vraag heeft beantwoord is 1552 en er zijn 3 missings. De missings zullen worden verwijderd en ook worden antwoordcategorieën 13 en 14 verwijderd.

**Syntax:** \*missings verwijderen.

```
FILTER OFF.
USE ALL.
SELECT IF (NOT (nettocat=0)).
EXECUTE.
```

**3.3 Uiteindelijke variabele**

De verdeling van de antwoordcategorieën is na de bewerking hetzelfde. De meeste respondenten geven aan tussen de 2001-2500 euro te verdienen.

**nettocat Personal net monthly income in categories**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 EUR 500 of minder	31	2,5	2,5	2,5
	2 EUR 501 t/m EUR 1000	98	7,8	7,8	10,2
	3 EUR 1001 t/m EUR 1500	159	12,6	12,6	22,9
	4 EUR 1501 t/m EUR 2000	287	22,8	22,8	45,7
	5 EUR 2001 t/m EUR 2500	289	23,0	23,0	68,6
	6 EUR 2501 t/m EUR 3000	189	15,0	15,0	83,6
	7 EUR 3001 t/m EUR 3500	95	7,5	7,5	91,2
	8 EUR 3501 t/m EUR 4000	58	4,6	4,6	95,8
	9 EUR 4001 t/m EUR 4500	24	1,9	1,9	97,7
	10 EUR 4501 t/m EUR 5000	8	,6	,6	98,3
	11 EUR 5001 t/m EUR 7500	14	1,1	1,1	99,4
	12 Meer dan EUR 7500	7	,6	,6	100,0
	Total	1259	100,0	100,0	

**Syntax:** \*beschrijvende stat. inkomen.

```
DESCRIPTIVES VARIABLES=nettocat
/STATISTICS=MEAN STDDEV VARIANCE MIN MAX KURTOSIS SKEWNESS.
```

**Variabele 'Fysiek zwaar werk'****4.1 Oorspronkelijke variabele**

De tweede mediator geeft de mate van fysiek zwaar werk van de respondent aan en heeft het label 'Mijn werk (is/was) fysiek veeleisend'. De variabele heeft de volgende antwoordcategorieën: 1 = "volledig mee oneens", 2 = "mee oneens", 3 = "mee eens", 4 = "volledig mee eens". De meeste respondenten geven aan het niet met de stelling eens te zijn, wat te zien is in de frequentietabel hieronder.

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std.		Skewness		Kurtosis	
					Deviation	Variance	Statistic	Std. Error	Statistic	Std. Error
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic
fysiekgzwaar My job [is/was] physically demanding.	1365	1	4	2,05	,859	,737	,409	,066	-,580	,132
Valid N (listwise)	1365									

**fysiekgzwaar My job [is/was] physically demanding.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 disagree entirely	400	25,7	29,3	29,3
	2 disagree	572	36,8	41,9	71,2
	3 agree	321	20,6	23,5	94,7
	4 agree entirely	72	4,6	5,3	100,0
	Total	1365	87,8	100,0	
Missing	System	190	12,2		
Total		1555	100,0		

**Syntax: \*beschrijvende statistieken fysiek zwaar werk.**

```
DESCRIPTIVES VARIABLES=fysiekgzwaar.
  /STATISTICS=MEAN STDDEV VARIANCE MIN MAX KURTOSIS SKEWNESS.
```

**Syntax: \*frequencies fysiek zwaar werk.**

```
FREQUENCIES VARIABLES=fysiekgzwaar.
  /NTILES=4
  /NTILES=10
  /STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT
  /HISTOGRAM
  /ORDER=ANALYSIS.
```

**4.2 Bewerking variabele**

Van de variabele zijn de 190 missings verwijderd.

**Syntax: \*verwijderen missings.**

```
FILTER OFF.
USE ALL.
SELECT IF (NOT (cw20m427=0)).
EXECUTE.
```

**4.3 Uiteindelijke variabele**

De verdeling van de antwoorden van de respondenten ziet er na het verwijderen van de missings nog hetzelfde uit.

**fysiekgwaar My job [is/was] physically demanding.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 disagree entirely	371	29,5	29,5	29,5
	2 disagree	530	42,1	42,1	71,6
	3 agree	294	23,4	23,4	94,9
	4 agree entirely	64	5,1	5,1	100,0
	Total	1259	100,0	100,0	

**Syntax:** \*frequencies fysiek zwaar werk.

DESCRIPTIVES VARIABLES=fysiekgwaar.

/STATISTICS=MEAN STDDEV VARIANCE MIN MAX KURTOSIS SKEWNESS.

## 5. Controlevariabelen

### Variabele 'Geslacht'

#### 5.1 Oorspronkelijke variabele

De eerste variabele waarvoor wordt gecontroleerd is geslacht. Hierbij geldt: 1 = "man", 2 = "vrouw". Er zitten ongeveer evenveel mannen (797) als vrouwen (758) in de dataset.

#### Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
							Statistic	Std. Error	Statistic	Std. Error
geslacht Gender	1555	1	2	1,49	,500	,250	,050	,062	-2,000	,124
Valid N (listwise)	1555									

#### geslacht Gender

		Frequency	Percent	Valid Percent	Cumulative
					Percent
Valid	1 Man	797	51,3	51,3	51,3
	2 Vrouw	758	48,7	48,7	100,0
	Total	1555	100,0	100,0	

#### Syntax: \*beschrijvende statistieken geslacht.

```
DESCRIPTIVES VARIABLES=geslacht
  /STATISTICS=MEAN STDDEV VARIANCE MIN MAX KURTOSIS SKEWNESS.
```

#### Syntax: \*frequencies geslacht.

```
FREQUENCIES VARIABLES=geslacht
  /NTILES=4
  /NTILES=10
  /STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT
  /HISTOGRAM
  /ORDER=ANALYSIS.
```

#### 5.2 Bewerking variabele

Van de variabele geslacht wordt een dummy gemaakt. Hierbij geldt: 0 = "man" en 1 = "vrouw".

#### Syntax: \*bewerken variabele geslacht.

```
RECODE geslacht (1=0) (2=1) INTO geslachtdummy.
EXECUTE.
```

**Variabele 'Leeftijd'****6.1 Oorspronkelijke variabele**

De tweede variabele waarvoor wordt gecontroleerd is leeftijd. De respondent kan zijn / haar leeftijd invullen. De gemiddelde leeftijd is 53, wat te vinden is in de descriptives hieronder. De variabele is redelijk normaal verdeeld.

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
							Statistic	Std. Error	Statistic	Std. Error
leeftijd Age of the household member	1555	40	69	53,27	7,393	54,654	-,082	,062	-1,029	,124
Valid N (listwise)	1555									

**leeftijd Age of the household member**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	40	46	3,0	3,0	3,0
	41	58	3,7	3,7	6,7
	42	51	3,3	3,3	10,0
	43	45	2,9	2,9	12,9
	44	35	2,3	2,3	15,1
	45	53	3,4	3,4	18,5
	46	68	4,4	4,4	22,9
	47	59	3,8	3,8	26,7
	48	56	3,6	3,6	30,3
	49	43	2,8	2,8	33,1
	50	65	4,2	4,2	37,2
	51	56	3,6	3,6	40,8
	52	62	4,0	4,0	44,8
	53	62	4,0	4,0	48,8
	54	67	4,3	4,3	53,1
	55	69	4,4	4,4	57,6
	56	78	5,0	5,0	62,6
	57	79	5,1	5,1	67,7

58	67	4,3	4,3	72,0
59	65	4,2	4,2	76,1
60	68	4,4	4,4	80,5
61	60	3,9	3,9	84,4
62	57	3,7	3,7	88,0
63	56	3,6	3,6	91,6
64	45	2,9	2,9	94,5
65	29	1,9	1,9	96,4
66	37	2,4	2,4	98,8
67	10	,6	,6	99,4
68	8	,5	,5	99,9
69	1	,1	,1	100,0
Total	1555	100,0	100,0	

**Syntax: \*beschrijvende statistieken leeftijd.**

```
DESCRIPTIVES VARIABLES=leeftijd
  /STATISTICS=MEAN STDDEV VARIANCE MIN MAX KURTOSIS SKEWNESS.
```

**Syntax: \*frequenties leeftijd.**

```
FREQUENCIES VARIABLES=leeftijd
  /NTILES=4
  /NTILES=10
  /STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT
  /HISTOGRAM
  /ORDER=ANALYSIS.
```

## 6.2 Bewerking variabele

Van de variabele leeftijd worden leeftijdscohorten gemaakt. Van de leeftijdscategorie 40-45 wordt leeftijdscohort 1 gemaakt, van de leeftijdscategorie 46-50 leeftijdscohort 2, van de leeftijdscategorie 51-55 leeftijdscohort 3, van de leeftijdscategorie 56-60 leeftijdscohort 4, van de leeftijdscategorie 61-65 leeftijdscohort 5 en van de leeftijdscategorie 66-69 leeftijdscohort 6.

**Syntax: \*bewerken variabele leeftijd.**

```
RECODE leeftijd (40 thru 45=1) (46 thru 50=2) (51 thru 55=3) (56 thru 60=4) (61 thru 65=5) (66 thru 69=6)
  INTO leeftijdcohort.
EXECUTE.
```

## 6.3 Uiteindelijke variabele

Na de bewerking zien de statistieken er nog hetzelfde uit. De meeste respondenten zitten in leeftijdscohort 4 (56-60 jaar).

		leeftijdcohort			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	1,00	218	17,3	17,3	17,3
	2,00	237	18,8	18,8	36,1
	3,00	257	20,4	20,4	56,6
	4,00	296	23,5	23,5	80,1

5,00	203	16,1	16,1	96,2
6,00	48	3,8	3,8	100,0
Total	1259	100,0	100,0	

**Syntax:** \*frequencies leeftijd.

```
FREQUENCIES VARIABLES=leeftijdcohort
/NTILES=4
/NTILES=10
/STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM MEAN MEDIAN MODE
/ORDER=ANALYSIS.
```

**Variabele 'Werksector'****7.1 Oorspronkelijke variabele**

De derde variabele waarvoor wordt gecontroleerd is de werksector waar de respondent in werkt. De variabele heeft de volgende antwoordcategorieën: 1 = "landbouw, jacht, bosbouw, visserij", 2 = "winning van delfstoffen", 3 = "industrie", 4 = "productie en distributie van en handel in elektriciteit, aardgas, stoom en water", 5 = "bouwnijverheid", 6 = "handel (incl. reparatie van consumentenartikelen)", 7 = "horeca", 8 = "vervoer, opslag en communicatie", 9 = "financiële instellingen", 10 = "zakelijke dienstverlening (incl. onroerend goed, verhuur van roerende goederen)", 11 = "overheidsdiensten, openbaar bestuur en verplichte sociale verzekeringen", 12 = "onderwijs", 13 = "gezondheids- en welzijnszorg", 14 = "milieudienstverlening, cultuur, recreatie en overige dienstverlening", 15 = "overig". De meeste respondenten werken in de gezondheids- en welzijnszorg. Er zijn 189 missings bij deze variabele.

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
							Statistic	Std. Error	Statistic	Std. Error
Werksector [In what sector do you work?/In what sector did you work in your last job?]	1366	1	15	9,78	3,936	15,489	-,539	,066	-,901	,132
Valid N (listwise)	1366									

**werksector [In what sector do you work?/In what sector did you work in your last job?]**

		Frequency	Percent	Valid Percent	Cumulative
					Percent
Valid	1 agriculture, forestry, fishery, hunting	25	1,6	1,8	1,8
	2 mining	1	,1	,1	1,9
	3 industrial production	148	9,5	10,8	12,7
	4 utilities production, distribution and/or trade (electricity, natural gas, steam)	9	,6	,7	13,4
	5 construction	59	3,8	4,3	17,7
	6 retail trade (including repairs of consumer goods)	123	7,9	9,0	26,7
	7 catering	36	2,3	2,6	29,4
	8 transport, storage and communication	78	5,0	5,7	35,1
	9 financial	53	3,4	3,9	38,9



10 business services (including real estate, rental)	94	6,0	6,9	45,8
11 government services, public administration and mandatory social insurances	155	10,0	11,3	57,2
12 education	116	7,5	8,5	65,7
13 healthcare and welfare	289	18,6	21,2	86,8
14 environmental services, culture, recreation and other services	34	2,2	2,5	89,3
15 other	146	9,4	10,7	100,0
Total	1366	87,8	100,0	
Missing System	189	12,2		
Total	1555	100,0		

**Syntax: \*beschrijvende statistieken werksector.**

```
DESCRIPTIVES VARIABLES=werksector
  /STATISTICS=MEAN STDDEV VARIANCE MIN MAX KURTOSIS SKEWNESS.
```

**Syntax: \*frequenties werksector.**

```
FREQUENCIES VARIABLES=werksector
  /NTILES=4
  /NTILES=10
  /STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT
  /HISTOGRAM
  /ORDER=ANALYSIS.
```

**7.2 Bewerking variabele**

Van de variabele worden dummy's gemaakt. Van de antwoordcategorieën 1, 2, 3 en 5 wordt de dummy gemaakt die aangeeft of de respondent in de sector 'industrie en bouw' werkt. Van de antwoordcategorieën 4, 6, 7, 8, 9 en 10 wordt een dummy gemaakt die aangeeft of de respondent in de sector 'diensten en handel' werkt. Van de antwoordcategorieën 11, 12, 13 en 14 wordt een dummy gemaakt die aangeeft of de respondent in de '(semi) publieke sector' werkt. Antwoordcategorie 15 (overig) zal de referentiecategorie worden. Dit is te zien in de syntax hieronder. Ook worden de missings verwijderd.

**Syntax: \*dummyindustriebouw.**

```
RECODE cw20m402 (1=1) (2=1) (3=1) (4=0) (5=1) (6=0) (7=0) (8=0) (9=0) (10=0) (11=0) (12=0) (13=0)
  (14=0) (15=0) (SYSMIS=SYSMIS) INTO dummyindustriebouw.
EXECUTE.
```

**Syntax: \*dummydienstenhandel.**

```
RECODE cw20m402 (1=0) (2=0) (3=0) (4=1) (5=0) (6=1) (7=1) (8=1) (9=1) (10=1) (11=0) (12=0) (13=0)
  (14=0) (15=0) (SYSMIS=SYSMIS) INTO dummydienstenhandel.
EXECUTE.
```

**Syntax: \*dummysemipubliek.**

```
RECODE cw20m402 (1=0) (2=0) (3=0) (4=0) (5=0) (6=0) (7=0) (8=0) (9=0) (10=0) (11=1) (12=1) (13=1)
  (14=1) (15=0) (SYSMIS=SYSMIS) INTO dummysemipubliek.
EXECUTE.
```

**7.3 Uiteindelijke variabele**

De antwoordcategorieën zijn samengevoegd tot verschillende dummy's. Hieronder volgt de frequenties. Er is te zien dat de meeste respondenten in de (semi)publieke sector werken.

**dummyindustrialbouw**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1043	82,8	82,8	82,8
	1,00	216	17,2	17,2	100,0
	Total	1259	100,0	100,0	

**dummydienstenhandel**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	898	71,3	71,3	71,3
	1,00	361	28,7	28,7	100,0
	Total	1259	100,0	100,0	

**dummysemipubliek**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	708	56,2	56,2	56,2
	1,00	551	43,8	43,8	100,0
	Total	1259	100,0	100,0	

**Syntax:** \*frequencies dummy's werksector.

```
FREQUENCIES VARIABLES=dummyindustrialbouw dienstenhandel semipubliek
  /STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM MEAN MEDIAN MODE
  /ORDER=ANALYSIS.
```

## Bijlage 2 – Statistische analyses

In deze bijlage zijn de statistische analyses van dit onderzoek weergegeven. In het eerste onderdeel zijn de bivariate statistieken van de variabelen gepresenteerd. In het tweede onderdeel volgen de modelschattingen.

### Onderdeel 1: Bivariate statistieken

#### Pearson Correlation

De correlaties tussen de continue variabelen is berekend aan de hand van een 'Pearson Correlation'. Dit geldt voor de variabelen: opleidingsniveau, inkomen, leeftijd en geslacht. De interpretatie van de Pearson correlaties is te vinden in de Resultatenparagraaf. Voor de samenhang tussen geslacht en de continue variabelen is ook een t-toets voor het verschil in gemiddelden gedaan. Dit is te vinden aan het einde van bijlage 2.

		Correlations			
		oplcat Level of education in CBS (Statistics Netherlands) categories	nettocat Personal net monthly income in categories	leeftijd Age of the household member	geslachtdummy
oplcat Level of education in CBS (Statistics Netherlands) categories	Pearson Correlation	1	,379**	-,045	-,019
	Sig. (2-tailed)		,000	,108	,506
	N	1259	1259	1259	1259
nettocat Personal net monthly income in categories	Pearson Correlation	,379**	1	-,013	-,378**
	Sig. (2-tailed)	,000		,636	,000
	N	1259	1259	1259	1259
leeftijd Age of the household member	Pearson Correlation	-,045	-,013	1	-,030
	Sig. (2-tailed)	,108	,636		,290
	N	1259	1259	1259	1259
geslachtdummy	Pearson Correlation	-,019	-,378**	-,030	1
	Sig. (2-tailed)	,506	,000	,290	
	N	1259	1259	1259	1259

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

**Syntax:** \*pearson correlatie continue variabelen.

```
CORRELATIONS
/VARIABLES=oplcat nettocat leeftijd geslachtdummy
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

### De bivariate samenhang uit een $\sqrt{R}$ Squared uit ANOVA

Voor de correlatie tussen een continue variabele en een categorische variabele wordt er gebruik gemaakt van een eenweg anova.

#### Variabele 'Boosheid' en 'Opleidingsniveau'

Uit de eenweg anova blijkt dat de groepsgemiddelden van de groepen van opleidingsniveau significant van elkaar verschillen in de mate waarop zij boos zijn. Ook blijkt dat de correlatie tussen de variabelen gelijk is aan -0.298. In de tabel descriptives is te zien dat de correlatie negatief is, omdat de groepsgemiddelden over het algemeen afnemen naarmate het opleidingsniveau hoger wordt.

#### Descriptive Statistics

Dependent Variable: v3d

oplcat Level of education in  
CBS (Statistics Netherlands)

categories	Mean	Std. Deviation	N
1 basisonderwijs	3,15	1,019	34
2 vmbo	2,80	1,200	193
3 havo/vwo	2,26	1,148	115
4 mbo	2,61	1,182	379
5 hbo	2,20	1,118	368
6 wo	1,75	,973	170
Total	2,39	1,184	1259

#### Tests of Between-Subjects Effects

Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	156,360 <sup>a</sup>	5	31,272	24,371	,000
Intercept	3998,635	1	3998,635	3116,230	,000
oplcat	156,360	5	31,272	24,371	,000
Error	1607,805	1253	1,283		
Total	8927,000	1259			
Corrected Total	1764,165	1258			

a. R Squared = ,089 (Adjusted R Squared = ,085)

**Syntax:** \*correlatie boosheid en opleidingsniveau.

```
UNIANOVA v3d BY oplcat
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /CRITERIA=ALPHA(0.05)
  /DESIGN=oplcat.
```

**Variabele 'Boosheid' en 'Inkomen'**

Uit de eenweg anova blijkt dat de groepsgemiddelden van de groepen van inkomen significant van elkaar verschillen in de mate waarop zij boos zijn. Ook blijkt dat de correlatie tussen de variabelen gelijk is aan -0.224. In de tabel descriptives is te zien dat de correlatie negatief is, omdat de groepsgemiddelden over het algemeen afnemen naarmate het inkomen hoger wordt.

**Descriptive Statistics**

Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

monthly income in categories	Mean	Std. Deviation	N
1 EUR 500 of minder	2,39	1,407	31
2 EUR 501 t/m EUR 1000	2,57	1,193	98
3 EUR 1001 t/m EUR 1500	2,63	1,156	159
4 EUR 1501 t/m EUR 2000	2,52	1,134	287
5 EUR 2001 t/m EUR 2500	2,52	1,190	289
6 EUR 2501 t/m EUR 3000	2,25	1,147	189
7 EUR 3001 t/m EUR 3500	1,94	1,099	95
8 EUR 3501 t/m EUR 4000	1,83	1,045	58
9 EUR 4001 t/m EUR 4500	1,92	1,213	24
10 EUR 4501 t/m EUR 5000	1,63	1,188	8
11 EUR 5001 t/m EUR 7500	1,50	1,092	14
12 Meer dan EUR 7500	1,71	1,254	7
Total	2,39	1,184	1259

**Tests of Between-Subjects Effects**

Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	88,109 <sup>a</sup>	11	8,010	5,959	,000
Intercept	1373,342	1	1373,342	1021,778	,000
nettocat	88,109	11	8,010	5,959	,000
Error	1676,056	1247	1,344		
Total	8927,000	1259			
Corrected Total	1764,165	1258			

a. R Squared = ,050 (Adjusted R Squared = ,042)

**Syntax: \*correlatie boosheid en inkomen.**

```
UNIANOVA v3d BY nettocat
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/PRINT=DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESIGN=nettocat.
```

**Variabele 'Boosheid' en 'Leeftijd'**

Uit de eenweg anova blijkt dat de gemiddelden van leeftijd significant van elkaar verschillen in de mate waarop zij boos zijn. Ook blijkt dat de correlatie tussen de variabelen gelijk is aan 0,202.

**Descriptive Statistics**

Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

leeftijd Age of the household member	Mean	Std. Deviation	N
40	2,36	1,055	33
41	2,04	1,065	45
42	2,49	1,193	37
43	2,38	1,184	39
44	2,48	,963	25
45	2,33	1,132	39
46	2,45	1,220	56
47	2,70	1,173	44
48	2,70	1,121	47
49	2,25	1,079	36
50	2,13	1,117	54
51	2,37	1,183	51
52	2,43	1,156	47
53	2,42	1,180	50
54	2,43	1,221	51
55	2,41	1,140	58
56	2,58	1,241	66
57	2,66	1,228	65
58	2,27	1,258	56
59	2,51	1,265	53
60	2,45	1,143	56
61	2,80	1,236	45
62	2,27	1,216	48
63	2,41	1,223	49
64	1,97	1,038	39
65	1,91	1,109	22
66	1,97	1,273	30
67	1,30	,949	10
68	1,57	1,134	7
69	1,00	.	1
Total	2,39	1,184	1259

### Tests of Between-Subjects Effects

Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	72,565 <sup>a</sup>	29	2,502	1,818	,005
Intercept	2472,939	1	2472,939	1796,667	,000
leeftijd	72,565	29	2,502	1,818	,005
Error	1691,600	1229	1,376		
Total	8927,000	1259			
Corrected Total	1764,165	1258			

a. R Squared = ,041 (Adjusted R Squared = ,019)

**Syntax:** \*correlatie boosheid en leeftijd.

```
UNIANOVA v3d BY leeftijd
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/PRINT=DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESIGN=leeftijd.
```

**Variabele 'Inkomen' en 'Werksector'**

Uit de eenweg anova blijkt dat de groepsgemiddelden van de groepen van werksector significant van elkaar verschillen in het inkomen dat ze verdienen. Er is te zien dat de groep 'mijnbouw' het hoogste scoort op inkomen. Echter is er maar 1 respondent die in de mijnbouw werkt, waardoor dit niet representatief is voor deze sector. Wanneer er naar de overige werksectoren wordt gekeken is te zien dat het meest wordt verdiend in werksector 4 (distributie/handel) en werksector 9 (financiële sector). Verder is hier niet naar de VR Squared uit ANOVA gekeken, omdat deze waarde hier geen betekenis heeft. Dit geldt ook voor de verdere associaties waar 'Werksector' onderdeel van uit maakt.

**Descriptive Statistics**

Dependent Variable: nettocat Personal net monthly income in categories

Mean	Std. Deviation	N	
cw20m402 [In what sector do you work?/In what sector did you work in your last job?]			
1 agriculture, forestry, fishery, hunting	4,60	2,010	20
2 mining	9,00	.	1
3 industrial production	5,77	2,066	137
4 utilities production, distribution and/or trade (electricity, natural gas, steam)	6,00	1,414	9
5 construction	4,88	1,687	58
6 retail trade (including repairs of consumer goods)	4,11	1,777	114
7 catering	3,91	1,692	32
8 transport, storage and communication	4,88	1,778	69
9 financial	5,90	2,347	49
10 business services (including real estate, rental)	5,64	1,852	88
11 government services, public administration and mandatory social insurances	5,74	1,728	147
12 education	4,89	1,455	108
13 healthcare and welfare	4,04	1,738	268
14 environmental services, culture, recreation and other services	4,54	1,875	28
15 other	4,34	1,948	131
Total	4,84	1,952	1259



### Tests of Between-Subjects Effects

Dependent Variable: nettocat Personal net monthly income in categories

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	674,150 <sup>a</sup>	14	48,154	14,551	,000
Intercept	4583,511	1	4583,511	1385,042	,000
cw20m402	674,150	14	48,154	14,551	,000
Error	4116,761	1244	3,309		
Total	34288,000	1259			
Corrected Total	4790,910	1258			

a. R Squared = ,141 (Adjusted R Squared = ,131)

**Syntax:** \*correlatie werksector en inkomen.

```
UNIANOVA nettocat BY cw20m402
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/PRINT=DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESIGN=cw20m402.
```

**Variabele 'Fysiek zwaar werk' en 'Opleidingsniveau'**

Uit de eenweg anova blijkt dat de groepsgemiddelden van de groepen van opleidingsniveau significant van elkaar verschillen in de mate waarop het werk als fysiek zwaar wordt ervaren. Ook blijkt dat de correlatie tussen de variabelen gelijk is aan -0.286. In de tabel descriptives is te zien dat de correlatie negatief is, omdat de groepsgemiddelden over het algemeen afnemen naarmate het opleidingsniveau hoger wordt.

**Descriptive Statistics**

Dependent Variable: cw20m427 My job [is/was] physically demanding.

oplcats Level of education in  
CBS (Statistics Netherlands)  
categories

	Mean	Std. Deviation	N
1 basisonderwijs	2,44	,991	34
2 vmbo	2,39	,853	193
3 havo/vwo	1,92	,850	115
4 mbo	2,21	,834	379
5 hbo	1,83	,801	368
6 wo	1,72	,732	170
Total	2,04	,855	1259

**Tests of Between-Subjects Effects**

Dependent Variable: cw20m427 My job [is/was] physically demanding.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	75,286 <sup>a</sup>	5	15,057	22,363	,000
Intercept	2870,949	1	2870,949	4263,982	,000
oplcats	75,286	5	15,057	22,363	,000
Error	843,648	1253	,673		
Total	6161,000	1259			
Corrected Total	918,934	1258			

a. R Squared = ,082 (Adjusted R Squared = ,078)

**Syntax:** \*correlatie fysiek zwaar en opleidingsniveau.

```
UNIANOVA cw20m427 BY oplcats
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/PRINT=DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESIGN=oplcats.
```

**Variabele 'Fysiek zwaar werk' en 'Inkomen'**

Uit de eenweg anova blijkt dat de groepsgemiddelden van de groepen van inkomen significant van elkaar verschillen in de mate waarop zij het werk als fysiek zwaar ervaren. Ook blijkt dat de correlatie tussen de variabelen gelijk is aan -0.311. In de tabel descriptives is te zien dat de correlatie negatief is, omdat de groepsgemiddelden over het algemeen afnemen naarmate het inkomen hoger wordt. Echter, er is een 'piek' te zien van het gemiddelde van de inkomenscategorie 5001 t/m 7500. De N is klein, wat een reden kan zijn voor het vinden het 'hoge' gemiddelde.

**Descriptive Statistics**

Dependent Variable: cw20m427 My job [is/was] physically demanding.

netto cat Personal net monthly income in categories	Mean	Std. Deviation	N
1 EUR 500 of minder	2,03	,875	31
2 EUR 501 t/m EUR 1000	2,37	,804	98
3 EUR 1001 t/m EUR 1500	2,43	,830	159
4 EUR 1501 t/m EUR 2000	2,20	,871	287
5 EUR 2001 t/m EUR 2500	2,01	,846	289
6 EUR 2501 t/m EUR 3000	1,79	,747	189
7 EUR 3001 t/m EUR 3500	1,63	,730	95
8 EUR 3501 t/m EUR 4000	1,67	,758	58
9 EUR 4001 t/m EUR 4500	1,50	,590	24
10 EUR 4501 t/m EUR 5000	1,50	,535	8
11 EUR 5001 t/m EUR 7500	2,07	,917	14
12 Meer dan EUR 7500	1,43	,787	7
Total	2,04	,855	1259

**Tests of Between-Subjects Effects**

Dependent Variable: cw20m427 My job [is/was] physically demanding.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	88,737 <sup>a</sup>	11	8,067	12,117	,000
Intercept	1089,997	1	1089,997	1637,234	,000
netto cat	88,737	11	8,067	12,117	,000
Error	830,197	1247	,666		
Total	6161,000	1259			
Corrected Total	918,934	1258			

a. R Squared = ,097 (Adjusted R Squared = ,089)

**Syntax:** \*correlatie fysiek zwaar en inkomen.

```
UNIANOVA cw20m427 BY netto cat
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/PRINT=DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESIGN=netto cat.
```

**Variabele 'Fysiek zwaar werk' en 'Leeftijd'**

Uit de eenweg anova blijkt dat de gemiddelden van leeftijd niet significant van elkaar verschillen in de mate waarop zij het werk als fysiek zwaar ervaren. Ook blijkt dat de correlatie tussen de variabelen gelijk is aan 0,126.

**Descriptive Statistics**

Dependent Variable: cw20m427 My job [is/was] physically demanding.

leeftijd Age of the household member	Mean	Std. Deviation	N
40	2,12	,740	33
41	2,00	,769	45
42	2,11	,994	37
43	2,21	,801	39
44	1,88	1,013	25
45	1,97	,707	39
46	2,09	,745	56
47	2,20	,978	44
48	1,96	,690	47
49	1,92	1,025	36
50	2,00	,890	54
51	2,10	,900	51
52	1,91	,830	47
53	2,02	,892	50
54	2,16	,946	51
55	2,03	,878	58
56	1,97	,744	66
57	2,14	,827	65
58	2,05	,942	56
59	2,09	,904	53
60	2,02	,842	56
61	2,18	,806	45
62	2,21	,798	48
63	1,96	,865	49
64	1,79	,833	39
65	1,82	,958	22
66	1,87	,937	30
67	2,10	,994	10
68	2,14	,378	7
69	2,00	.	1
Total	2,04	,855	1259

### Tests of Between-Subjects Effects

Dependent Variable: cw20m427 My job [is/was] physically demanding.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	14,364 <sup>a</sup>	29	,495	,673	,906
Intercept	1988,571	1	1988,571	2701,784	,000
leeftijd	14,364	29	,495	,673	,906
Error	904,570	1229	,736		
Total	6161,000	1259			
Corrected Total	918,934	1258			

a. R Squared = ,016 (Adjusted R Squared = -,008)

**Syntax:** \*correlatie fysiek zwaar en leeftijd.

UNIANOVA cw20m427 BY leeftijd

/METHOD=SSTYPE(3)

/INTERCEPT=INCLUDE

/PRINT=DESCRIPTIVE

/CRITERIA=ALPHA(.05)

/DESIGN=leeftijd.

**Variabele 'Opleidingsniveau' en 'Werksector'**

Uit de eenweg anova blijkt dat de gemiddelden van werksector significant van elkaar verschillen in het opleidingsniveau. Er is te zien dat de respondent die in de mijnbouw werkt het hoogste scoort. Deze persoon heeft het WO gedaan (antwoordcategorie 5) en verdient ook een hoog inkomen (gebleken uit de samenhang 'Inkomen' en 'Werksector'). Wanneer er naar de andere variabelen wordt gekeken, is te zien dat de respondenten in werksector 10 (zakelijke dienstverlening), werksector 12 (onderwijs) en werksector 4 (productie en distributie/handel in elektriciteit etc.) gemiddeld ook een hoog opleidingsniveau hebben. Het laagste gemiddelde opleidingsniveau ligt in werksector 7 (de catering).

**Descriptive Statistics**

Dependent Variable: oplcat Level of education in CBS (Statistics

Netherlands) categories

cw20m402 [In what sector do you work?/In what sector did you work in your last job?]	Mean	Std. Deviation	N
1 agriculture, forestry, fishery, hunting	3,75	1,372	20
2 mining	5,00	.	1
3 industrial production	3,99	1,342	137
4 utilities production, distribution and/or trade (electricity, natural gas, steam	4,89	,601	9
5 construction	3,66	1,371	58
6 retail trade (including repairs of consumer goods)	3,53	1,131	114
7 catering	3,25	1,437	32
8 transport, storage and communication	3,55	1,399	69
9 financial	4,47	1,063	49
10 business services (including real estate, rental)	4,82	1,109	88
11 government services, public administration and mandatory social insurances	4,44	1,182	147
12 education	4,76	1,229	108
13 healthcare and welfare	4,11	1,206	268
14 environmental services, culture, recreation and other services	4,39	1,397	28
15 other	3,61	1,465	131
Total	4,08	1,332	1259

### Tests of Between-Subjects Effects

Dependent Variable: oplcat Level of education in CBS (Statistics Netherlands) categories

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	252,377 <sup>a</sup>	14	18,027	11,327	,000	,113
Intercept	2898,906	1	2898,906	1821,456	,000	,594
cw20m402	252,377	14	18,027	11,327	,000	,113
Error	1979,866	1244	1,592			
Total	23225,000	1259				
Corrected Total	2232,243	1258				

a. R Squared = ,113 (Adjusted R Squared = ,103)

**Syntax:** \*correlatie werksector en opleidingsniveau.

```
UNIANOVA oplcat BY cw20m402
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/PRINT=DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESIGN=cw20m402.
```

**Variabele 'Werksector' en 'Leeftijd'**

Uit de eenweg anova blijkt dat de gemiddelden van werksector significant van elkaar verschillen in hun leeftijd. De oudste gemiddelde leeftijd ligt in de werksector 'Onderwijs'. De persoon uit de mijnbouw is één van de jongste respondenten uit de steekproef, namelijk 41 jaar. De leeftijden van de meeste sectoren liggen gemiddeld rond de 51-55 jaar.

**Descriptive Statistics**

Dependent Variable: leeftijd Age of the household member

job?]	Mean	Std. Deviation	N
1 agriculture, forestry, fishery, hunting	54,15	7,286	20
2 mining	41,00	.	1
3 industrial production	53,22	6,647	137
4 utilities production, distribution and/or trade (electricity, natural gas, steam	51,44	8,472	9
5 construction	51,74	7,290	58
6 retail trade (including repairs of consumer goods)	52,48	6,512	114
7 catering	53,31	8,110	32
8 transport, storage and communication	52,23	7,009	69
9 financial	52,37	8,343	49
10 business services (including real estate, rental)	52,17	8,364	88
11 government services, public administration and mandatory social insurances	54,22	7,231	147
12 education	55,31	8,023	108
13 healthcare and welfare	54,47	7,103	268
14 environmental services, culture, recreation and other services	52,68	7,469	28
15 other	53,22	7,128	131
Total	53,47	7,346	1259



### Tests of Between-Subjects Effects

Dependent Variable: leeftijd Age of the household member

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1555,014 <sup>a</sup>	14	111,072	2,083	,011
Intercept	460419,352	1	460419,352	8635,993	,000
cw20m402	1555,014	14	111,072	2,083	,011
Error	66322,619	1244	53,314		
Total	3667439,000	1259			
Corrected Total	67877,633	1258			

a. R Squared = ,023 (Adjusted R Squared = ,012)

**Syntax: \*correlatie werksector en leeftijd.**

```
UNIANOVA leeftijd BY cw20m402
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/PRINT=DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESIGN=cw20m402.
```

**Cramer's V**

Voor de samenhang van twee categorische variabelen wordt gekeken naar de Cramer's V. De waarde ligt tussen 0 (geen samenhang) en 1 (perfecte samenhang).

**Variabele 'Boosheid' en 'Geslacht'**

De Cramer's V voor deze associatie is 0,116 en de associatie is significant. De samenhang tussen de variabelen is zwak. In de tabel hieronder is te zien dat de meeste mannen en de meeste vrouwen niet boos zijn dat ze langer moeten doorwerken.

**v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan \* geslachtdummy Crosstabulation**

Count		geslachtdummy		Total
		,00	1,00	
v3d Als u denkt aan werken	1 Niet	231	171	402
tot uw AOW-leeftijd, in	2 In geringe mate	131	170	301
hoeverre bent u boos dat u	3 In redelijke mate	106	119	225
zo lan	4 In hoge mate	183	148	331
Total		651	608	1259

**Symmetric Measures**

		Value	Approximate Significance
Nominal by Nominal	Phi	,116	,001
	Cramer's V	,116	,001
N of Valid Cases		1259	

**Syntax: \*boosheid en geslacht**

```
CROSSTABS
/TABLES=geslachtdummy BY v3d
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.
```

**Variabele 'Boosheid' en 'Fysiek zwaar werk'**

De Cramer's V voor deze associatie is 0,146 en de associatie is significant. De samenhang tussen de twee variabelen is zwak. Er is te zien dat de respondenten die het werk als (erg) fysiek zwaar ervaren, relatief hoger scoren op de mate van boosheid dan de respondenten die het werk niet als fysiek zwaar ervaren.

**v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan \* cw20m427 My job [is/was] physically demanding. Crosstabulation**

Count

		cw20m427 My job [is/was] physically demanding.				Total
		1 disagree entirely	2 disagree	3 agree	4 agree entirely	
v3d Als u denkt aan	1 Niet	159	168	65	10	402
werken tot uw AOW- leeftijd, in hoeverre	2 In geringe mate	89	137	68	7	301
bent u boos dat u zo lan	3 In redelijke mate	62	93	60	10	225
	4 In hoge mate	61	132	101	37	331
Total		371	530	294	64	1259

**Symmetric Measures**

		Value	Approximate Significance
Nominal by Nominal	Phi	,253	,000
	Cramer's V	,146	,000
N of Valid Cases		1259	

**Syntax: \*fysiek zwaar werk en boosheid.**

```
CROSSTABS
/TABLES=cw20m427 BY v3d
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.
```

**Variabele 'Fysiek zwaar werk' en 'Geslacht'**

De Cramer's V van deze associatie is 0,121 en de associatie is significant. De twee variabelen hangen zwak samen. De verdeling van de antwoorden van man en vrouw is ongeveer gelijk.

**cw20m427 My job [is/was] physically demanding. \* geslachtdummy**  
**Crosstabulation**

Count

		geslachtdummy		Total
		,00	1,00	
cw20m427 My job [is/was]	1 disagree entirely	190	181	371
physically demanding.	2 disagree	306	224	530
	3 agree	125	169	294
	4 agree entirely	30	34	64
<b>Total</b>		<b>651</b>	<b>608</b>	<b>1259</b>

**Symmetric Measures**

		Value	Approximate Significance
Nominal by Nominal	Phi	,121	,000
	Cramer's V	,121	,000
N of Valid Cases		1259	

**Syntax: \*fysiek zwaar werk en geslacht.**

```
CROSSTABS
/TABLES=geslachtdummy BY cw20m427
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.
```

**Variabele 'Werksector' en 'Geslacht'**

De Cramer's V van deze associatie is 0,416 en de associatie is significant. De twee variabelen hangen matig samen. Er is te zien dat veel vrouwelijke respondenten in de zorg werken en veel mannen in de industriële sector en de bouwnijverheid.

**cw20m402 [In what sector do you work?/In what sector did you work in your last job?] \* geslachtdummy Crosstabulation**

Count		geslachtdummy		Total
		,00	1,00	
cw20m402 [In what sector do you work?/In what sector did you work in your last job?]	1 agriculture, forestry, fishery, hunting	14	6	20
	2 mining	1	0	1
	3 industrial production	113	24	137
	4 utilities production, distribution and/or trade (electricity, natural gas, steam)	7	2	9
	5 construction	45	13	58
	6 retail trade (including repairs of consumer goods)	57	57	114
	7 catering	15	17	32
	8 transport, storage and communication	54	15	69
	9 financial	30	19	49
	10 business services (including real estate, rental)	58	30	88
	11 government services, public administration and mandatory social insurances	82	65	147
	12 education	42	66	108
	13 healthcare and welfare	54	214	268
	14 environmental services, culture, recreation and other services	13	15	28
	15 other	66	65	131
<b>Total</b>	<b>651</b>	<b>608</b>	<b>1259</b>	

**Symmetric Measures**

		Value	Approximate Significance
Nominal by Nominal	Phi	,416	,000
	Cramer's V	,416	,000

N of Valid Cases	1259
------------------	------

**Syntax:** \*werksector en geslacht.

```
CROSSTABS
```

```
/TABLES=geslachtdummy BY cw20m402
```

```
/FORMAT=AVALUE TABLES
```

```
/STATISTICS=CHISQ PHI
```

```
/CELLS=COUNT
```

```
/COUNT ROUND CELL.
```

**Variabele 'Werksector' en 'Fysiek zwaar werk'**

De Cramer's V van deze associatie is 0,221 en de associatie is significant. De twee variabelen hangen zwak samen. Er is te zien dat het overgrote deel van de respondenten in de gezondheidszorg het werk als fysiek veeleisend ervaart. De antwoorden van de respondenten in de bouwnijverheid (werksector 5) zijn meer gelijk verdeeld over de mate waarin het werk als fysiek veeleisend wordt ervaren. Respondenten die werken bij de overheid en de zakelijke dienstverlening ervaren over het algemeen het werk niet als fysiek veeleisend.

**cw20m402 [In what sector do you work?/In what sector did you work in your last job?] \***  
**cw20m427 My job [is/was] physically demanding. Crosstabulation**

Count

		cw20m427 My job [is/was] physically demanding.				Total
		1 disagree entirely	2 disagree	3 agree	4 agree entirely	
cw20m402 [In what sector do you work?/In what sector did you work in your last job?]	1 agriculture, forestry, fishery, hunting	6	7	5	2	
	2 mining	0	1	0	0	1
	3 industrial production	37	60	32	8	137
	4 utilities production, distribution and/or trade (electricity, natural gas, steam)	4	5	0	0	9
	5 construction	15	22	15	6	58
	6 retail trade (including repairs of consumer goods)	30	48	31	5	114
	7 catering	3	11	13	5	32
	8 transport, storage and communication	20	29	16	4	69
	9 financial	27	20	0	2	49
	10 business services (including real estate, rental)	50	31	6	1	88
	11 government services, public administration and mandatory social insurances	67	70	9	1	147
	12 education	23	63	21	1	108
	13 healthcare and welfare	49	96	102	21	268
	14 environmental services, culture, recreation and other services	6	13	7	2	28
	15 other	34	54	37	6	131
<b>Total</b>		<b>371</b>	<b>530</b>	<b>294</b>	<b>64</b>	<b>1259</b>

**Symmetric Measures**

		Value	Approximate Significance
Nominal by Nominal	Phi	,382	,000
	Cramer's V	,221	,000
N of Valid Cases		1259	

**Syntax:** \*werksector en fysiek zwaar.

```
CROSSTABS
/TABLES=cw20m427 BY cw20m402
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.
```



**Variabele 'Werksector' en 'Boosheid'**

De Cramer's V van deze associatie is 0,109 en de associatie is niet significant. De twee variabelen hangen zwak met elkaar samen. Er is te zien dat de respondenten in de sector 'Industrie en Bouw' (werksectoren 1, 2, 3 en 5) relatief in hoge mate boos zijn omtrent het langer doorwerken. Ook zijn de meeste respondenten in de gezondheids- en welvaartzorg in hoge mate boos. De zakelijke dienstverlening en overheidsdiensten scoren het hoogst op het niet boos zijn omtrent het langer moeten doorwerken.

**cw20m402 [In what sector do you work?/In what sector did you work in your last job?] \* v3d**  
**Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan**  
**Crosstabulation**

Count

		v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan				
		lan				
		1 Niet	2 In geringe mate	3 In redelijke mate	4 In hoge mate	Total
cw20m402 [In what sector do you work?/In what sector did you work in your last job?]	1 agriculture, forestry, fishery, hunting	4	7	1	8	20
	2 mining	0	0	1	0	1
	3 industrial production	38	28	30	41	137
	4 utilities production, distribution and/or trade (electricity, natural gas, steam)	1	3	2	3	9
	5 construction	20	11	9	18	58
	6 retail trade (including repairs of consumer goods)	34	26	20	34	114
	7 catering	11	8	4	9	32
	8 transport, storage and communication	19	15	11	24	69
	9 financial	18	14	12	5	49
	10 business services (including real estate, rental)	40	21	14	13	88
	11 government services, public administration and mandatory social insurances	51	36	24	36	147
	12 education	37	31	16	24	108
	13 healthcare and welfare	74	64	54	76	268
	14 environmental services, culture, recreation and other services	12	7	4	5	28
	15 other	43	30	23	35	131
<b>Total</b>		<b>402</b>	<b>301</b>	<b>225</b>	<b>331</b>	<b>1259</b>

**Symmetric Measures**

		Value	Approximate Significance
Nominal by Nominal	Phi	,189	,348
	Cramer's V	,109	,348
N of Valid Cases		1259	

**Syntax:** \*werksector en boosheid.

```
CROSSTABS
/TABLES=cw20m402 BY v3d
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.
```

### T-toets voor twee gemiddelden voor geslacht

Voor de samenhang tussen de continue variabele en geslacht is ook een t-toets voor twee gemiddelden uitgevoerd. Hierdoor kan er een betere interpretatie aan de samenhang worden gegeven. Uit de t-toets blijkt dat er een significant verschil zit in de gemiddelden van mannen en vrouwen voor de variabele inkomen ( $t(1257)=14,477, p<0,001$ ). Hierbij is te zien dat mannen over het algemeen 1,5 punt hoger scoren op het inkomen dan vrouwen en dus een hoger netto inkomen per maand hebben. Voor de variabelen opleiding en leeftijd is weinig verschil te zien in de gemiddelden. De verschillen zijn ook niet significant.

#### Group Statistics

	geslachtdummy	N	Mean	Std. Deviation	Std. Error Mean
oplcat Level of education in CBS (Statistics Netherlands) categories	man	651	4,11	1,351	,053
	vrouw	608	4,06	1,312	,053
nettocat Personal net monthly income in categories	man	651	5,55	1,887	,074
	vrouw	608	4,08	1,718	,070
leeftijd Age of the household member	man	651	53,68	7,422	,291
	vrouw	608	53,24	7,262	,295

#### Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
oplcat Level of education in CBS (Statistics Netherlands) categories	Equal variances assumed	,683	,409	,665	1257	,506	,050	,075	-,097	,197
	Equal variances not assumed			,666	1256,078	,506	,050	,075	-,097	,197
nettocat Personal net monthly income in categories	Equal variances assumed	3,787	,052	14,477	1257	,000	1,476	,102	1,276	1,676
	Equal variances not assumed			14,523	1256,180	,000	1,476	,102	1,276	1,675
leeftijd Age of the household member	Equal variances assumed	,384	,536	1,059	1257	,290	,439	,414	-,374	1,251
	Equal variances not assumed			1,060	1254,272	,290	,439	,414	-,374	1,251

**Syntax: \*t-toets geslacht en continue variabelen.**

```
T-TEST GROUPS=geslachtdummy (0 1)
/MISSING=ANALYSIS
/VARIABLES=oplcat nettocat leeftijd
/CRITERIA=CI(.95).
```

Onderdeel 2: Modelschattingen

In dit onderdeel worden de verschillende modellen weergegeven van de regressieanalyse. Per model zijn de regressiecoëfficiënten, de modelfit en de  $R^2_{\text{change}}$  waarden weergegeven.

**Schatting 1: regressieanalyse met de afhankelijke variabele 'Boosheid' en zonder model 4b**

Allereerst wordt de syntax en bijbehorende output van de lineaire regressieanalyse met als afhankelijke variabele boosheid weergegeven. Deze regressieanalyse is op twee manieren geschat. De eerste schatting zonder model 4b en de tweede schatting zonder model 4a. Hierdoor ontstaat er een hiërarchische regressieanalyse en wordt er betekenis gegeven aan de  $R^2_{\text{change}}$  waarden. Hieronder volgt de output en syntax van de regressieanalyse zonder model 4b en enkel de modellen 1, 2, 4a en 5.

**Model Summary<sup>e</sup>**

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	,069 <sup>a</sup>	,005	,001	1,184	,005	1,213	5	1253	,301
2	,270 <sup>b</sup>	,073	,068	1,143	,068	91,804	1	1252	,000
3	,293 <sup>c</sup>	,086	,081	1,135	,013	17,780	1	1251	,000
4	,331 <sup>d</sup>	,109	,104	1,121	,023	32,928	1	1250	,000

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	2,435	,132		18,410	,000		
	dummybouw	,156	,133	,050	1,177	,239	,444	2,252
	industrialhandel	-,058	,121	-,022	-,482	,630	,371	2,693
	semipubliek	-,010	,116	-,004	-,082	,934	,337	2,966
	geslachtdummy	,056	,071	,023	,779	,436	,878	1,139
	leeftijdcohort	-,026	,023	-,032	-1,129	,259	,978	1,023
2	(Constant)	3,359	,160		20,990	,000		
	dummybouw	,205	,128	,065	1,601	,110	,443	2,256
	industrialhandel	,022	,117	,009	,191	,848	,369	2,707
	semipubliek	,175	,113	,073	1,539	,124	,327	3,054
	geslachtdummy	,006	,069	,002	,083	,934	,873	1,145
	leeftijdcohort	-,040	,023	-,049	-1,782	,075	,974	1,027
	oplcats Level of education in CBS (Statistics Netherlands) categories	-,237	,025	-,267	-9,581	,000	,956	1,045

3	(Constant)	3,610	,170		21,266	,000		
	dummybouw	,249	,128	,079	1,948	,052	,440	2,271
	industriehandel	,039	,116	,015	,335	,738	,369	2,710
	semipubliek	,188	,113	,079	1,665	,096	,327	3,056
	geslachtdummy	-,107	,074	-,045	-1,450	,147	,759	1,318
	leeftijdcohort	-,041	,022	-,050	-1,823	,069	,974	1,027
	oplcat Level of education in CBS (Statistics Netherlands) categories	-,192	,027	-,216	-7,169	,000	,805	1,242
	netto Personal net monthly income in categories	-,082	,019	-,135	-4,217	,000	,713	1,403
4	(Constant)	2,931	,205		14,289	,000		
	dummybouw	,223	,126	,071	1,762	,078	,440	2,274
	industriehandel	,067	,115	,026	,583	,560	,368	2,715
	semipubliek	,167	,111	,070	1,496	,135	,327	3,060
	geslachtdummy	-,092	,073	-,039	-1,266	,206	,758	1,320
	leeftijdcohort	-,036	,022	-,044	-1,631	,103	,972	1,028
	oplcat Level of education in CBS (Statistics Netherlands) categories	-,167	,027	-,188	-6,235	,000	,784	1,276
	netto Personal net monthly income in categories	-,061	,020	-,100	-3,114	,002	,688	1,454
	cw20m427 My job [is/was] physically demanding.	,225	,039	,162	5,738	,000	,892	1,121

a. Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

**Syntax: \*schatting model met inkomen.**

REGRESSION

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE

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

/NOORIGIN

/DEPENDENT v3d

/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort

/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat

/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat netto

/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort netto

cw20m427 oplcat

/SCATTERPLOT=(\*ZRESID , \*ZPRED) .

**Schatting 2: regressieanalyse met de afhankelijke variabele 'Boosheid' en zonder model 4a**

Hieronder volgt de output en de syntax van de regressieanalyse zonder model 4a. Hierin zijn enkel de modellen 1, 2, 4b en 5 te zien.

**Model Summary<sup>e</sup>**

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	,069 <sup>a</sup>	,005	,001	1,184	,005	1,213	5	1253	,301
2	,270 <sup>b</sup>	,073	,068	1,143	,068	91,804	1	1252	,000
3	,320 <sup>c</sup>	,102	,097	1,125	,030	41,182	1	1251	,000
4	,331 <sup>d</sup>	,109	,104	1,121	,007	9,694	1	1250	,002

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	3,359	,160		20,990	,000	3,045	3,672
	dummybouw	,205	,128	,065	1,601	,110	-,046	,457
	industriehandel	,022	,117	,009	,191	,848	-,208	,252
	semipubliek	,175	,113	,073	1,539	,124	-,048	,397
	geslachtdummy	,006	,069	,002	,083	,934	-,130	,141
	leeftijdcohort	-,040	,023	-,049	-1,782	,075	-,085	,004
	oplcats Level of education in CBS (Statistics Netherlands) categories	-,237	,025	-,267	-9,581	,000	-,286	-,188
2	(Constant)	2,931	,205		14,289	,000	2,529	3,334
	dummybouw	,223	,126	,071	1,762	,078	-,025	,471
	industriehandel	,067	,115	,026	,583	,560	-,159	,293
	semipubliek	,167	,111	,070	1,496	,135	-,052	,385
	geslachtdummy	-,092	,073	-,039	-1,266	,206	-,234	,051
	leeftijdcohort	-,036	,022	-,044	-1,631	,103	-,080	,007
	oplcats Level of education in CBS (Statistics Netherlands) categories	-,167	,027	-,188	-6,235	,000	-,220	-,115
cw20m427 My job [is/was] physically demanding.	,225	,039	,162	5,738	,000	,148	,302	

nettocat Personal net monthly income in categories	-,061	,020	-,100	-3,114	,002	-,099	-,023
--	-------	------	-------	--------	------	-------	-------

a. Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

**Syntax: \*model met fysiek zwaar werk.**

REGRESSION

```

/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT v3d
/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort
/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat
/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat cw20m427
/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort nettocat
      cw20m427 oplcat
/SCATTERPLOT=(*ZRESID ,*ZPRED) .

```

De modellen 3a en 3b zijn apart geschat. Model 3a heeft als afhankelijke variabele inkomen en model 4b heeft als afhankelijke variabele de mate van fysiek zwaar werk. Hieronder volgt de syntax en output van beide modellen.

### Schatting 3: regressieanalyse met als afhankelijke variabele 'Inkomen'

Hieronder volgt de output en syntax van de analyse met als afhankelijke variabele 'Inkomen'. Dit is model 3a. In model 3a is getoetst of opleidingsniveau significant verschilt in het inkomen.

		Coefficients <sup>a</sup>						
Model		Unstandardized Coefficients		Standardized	t	Sig.	Collinearity Statistics	
		B	Std. Error	Coefficients Beta			Tolerance	VIF
1	(Constant)	3,066	,231		13,261	,000		
	dummybouw	,534	,185	,103	2,879	,004	,443	2,256
	industriehandel	,202	,169	,047	1,196	,232	,369	2,707
	semipubliek	,160	,164	,041	,976	,329	,327	3,054
	geslachtdummy	-1,371	,100	-,351	-13,750	,000	,873	1,145
	leeftijdcohort	-,008	,033	-,006	-,239	,811	,974	1,027
	oplcat Level of education in CBS (Statistics Netherlands) categories	,549	,036	,375	15,357	,000	,956	1,045

a. Dependent Variable: nettocat Personal net monthly income in categories

**Syntax: \*model3a.**

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT nettocat
/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat
/SCATTERPLOT=(*ZRESID ,*ZPRED).
```



**Schatting 4: regressieanalyse met als afhankelijke variabele 'Fysiek zwaar werk'**

Hieronder volgt de output en syntax van de analyse met als afhankelijke variabele 'Fysiek zwaar werk'. Dit is model 3b. In model 3b is getoetst of opleidingsniveau significant verschilt in de mate van fysiek zwaar werk.

Model		Coefficients <sup>a</sup>					Collinearity Statistics	
		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Tolerance	VIF
		B	Std. Error	Beta				
1	(Constant)	2,732	,115		23,685	,000		
	dummybouw	,068	,093	,030	,734	,463	,443	2,256
	industriehandel	-,144	,084	-,076	-1,709	,088	,369	2,707
	semipubliek	,079	,082	,046	,967	,334	,327	3,054
	geslachtdummy	,064	,050	,037	1,278	,202	,873	1,145
	leeftijdcohort	-,020	,016	-,034	-1,247	,213	,974	1,027
	oplcats Level of education in CBS (Statistics Netherlands) categories	-,162	,018	-,253	-9,109	,000	,956	1,045

a. Dependent Variable: cw20m427 My job [is/was] physically demanding.

**Syntax: \*model3b.**

```
*model3b.
```

```
REGRESSION
```

```
  /MISSING LISTWISE
```

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

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

```
  /NOORIGIN
```

```
  /DEPENDENT cw20m427
```

```
  /METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcats
```

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

**Betrouwbaarheidsintervallen**

Aan de hand van de betrouwbaarheidsintervallen is te zien of de afname van de helling van de onafhankelijke variabele (opleidingsniveau) ook significant is. Allereerst wordt er gekeken naar het betrouwbaarheidsinterval van opleidingsniveau in model 2 en 4a. In model 2 is het BHI -0,286;-0,188 en in model 4a -0,245;-0,139. De betrouwbaarheidsintervallen overlappen elkaar grotendeels.

		Coefficients <sup>a</sup>					95,0% Confidence Interval for	
		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	B	
Model	B	Std. Error	Beta					Lower Bound
1	(Constant)	3,359	,160		20,990	,000	3,045	3,672
	dummybouw	,205	,128	,065	1,601	,110	-,046	,457
	industriehandel	,022	,117	,009	,191	,848	-,208	,252
	semipubliek	,175	,113	,073	1,539	,124	-,048	,397
	geslachtdummy	,006	,069	,002	,083	,934	-,130	,141
	leeftijdcohort	-,040	,023	-,049	-1,782	,075	-,085	,004
	oplcats Level of education in CBS (Statistics Netherlands) categories	-,237	,025	-,267	-9,581	,000	-,286	-,188
2	(Constant)	3,610	,170		21,266	,000	3,277	3,943
	dummybouw	,249	,128	,079	1,948	,052	-,002	,500
	industriehandel	,039	,116	,015	,335	,738	-,190	,268
	semipubliek	,188	,113	,079	1,665	,096	-,033	,409
	geslachtdummy	-,107	,074	-,045	-1,450	,147	-,251	,038
	leeftijdcohort	-,041	,022	-,050	-1,823	,069	-,085	,003
	oplcats Level of education in CBS (Statistics Netherlands) categories	-,192	,027	-,216	-7,169	,000	-,245	-,139
nettocat Personal net monthly income in categories	-,082	,019	-,135	-4,217	,000	-,120	-,044	

a. Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

**Syntax: \*betrouwbaarheidsintervallen model 2 en 4a.**

```
REGRESSION
  /DESCRIPTIVES MEAN STDDEV CORR SIG N
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS CI(95) R ANOVA CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT v3d
  /METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat
  /METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat nettocat.
```

Vervolgens wordt er gekeken naar het betrouwbaarheidsinterval van opleidingsniveau in model 2 en model 4b. Het BHI van opleidingsniveau in model 2 is -0,286;-0,188 en in model 4b -0,246;-0,147.

		Coefficients <sup>a</sup>					95,0% Confidence Interval for	
		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	B	
Model	B	Std. Error	Beta					Lower Bound
1	(Constant)	3,359	,160		20,990	,000	3,045	3,672
	dummybouw	,205	,128	,065	1,601	,110	-,046	,457
	industriehandel	,022	,117	,009	,191	,848	-,208	,252
	semipubliek	,175	,113	,073	1,539	,124	-,048	,397
	geslachtdummy	,006	,069	,002	,083	,934	-,130	,141
	leeftijdcohort	-,040	,023	-,049	-1,782	,075	-,085	,004
	oplcats Level of education in CBS (Statistics Netherlands) categories	-,237	,025	-,267	-9,581	,000	-,286	-,188
2	(Constant)	2,682	,190		14,151	,000	2,310	3,054
	dummybouw	,189	,126	,060	1,493	,136	-,059	,437
	industriehandel	,058	,115	,022	,504	,615	-,168	,285
	semipubliek	,155	,112	,065	1,388	,165	-,064	,374
	geslachtdummy	-,010	,068	-,004	-,148	,883	-,143	,123
	leeftijdcohort	-,035	,022	-,043	-1,583	,114	-,079	,008
	oplcats Level of education in CBS (Statistics Netherlands) categories	-,197	,025	-,221	-7,827	,000	-,246	-,147
	cw20m427 My job [is/was] physically demanding.	,248	,039	,179	6,417	,000	,172	,323

a. Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

**Syntax: \*betrouwbaarheidsintervallen model 2 en 4b.**

REGRESSION

/DESCRIPTIVES MEAN STDDEV CORR SIG N

/MISSING LISTWISE

/STATISTICS COEFF OUTS CI(95) R ANOVA CHANGE

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

/NOORIGIN

/DEPENDENT v3d

/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcats

/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcats cw20m427.

Als laatst wordt er gekeken naar de betrouwbaarheidsintervallen van opleidingsniveau in model 2 en model 5 (wanneer beide mediators aan het model zijn toegevoegd). In model 2 is het BHI -0,286;-0,188 en in model 5 -0,220;-0,115. De betrouwbaarheidsintervallen overlappen deels.

		Coefficients <sup>a</sup>					95,0% Confidence Interval for	
		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	B	
Model	B	Std. Error	Beta					Lower Bound
1	(Constant)	3,359	,160		20,990	,000	3,045	3,672
	dummybouw	,205	,128	,065	1,601	,110	-,046	,457
	industriehandel	,022	,117	,009	,191	,848	-,208	,252
	semipubliek	,175	,113	,073	1,539	,124	-,048	,397
	geslachtdummy	,006	,069	,002	,083	,934	-,130	,141
	leeftijdcohort	-,040	,023	-,049	-1,782	,075	-,085	,004
	oplcats Level of education in CBS (Statistics Netherlands) categories	-,237	,025	-,267	-9,581	,000	-,286	-,188
2	(Constant)	2,931	,205		14,289	,000	2,529	3,334
	dummybouw	,223	,126	,071	1,762	,078	-,025	,471
	industriehandel	,067	,115	,026	,583	,560	-,159	,293
	semipubliek	,167	,111	,070	1,496	,135	-,052	,385
	geslachtdummy	-,092	,073	-,039	-1,266	,206	-,234	,051
	leeftijdcohort	-,036	,022	-,044	-1,631	,103	-,080	,007
	oplcats Level of education in CBS (Statistics Netherlands) categories	-,167	,027	-,188	-6,235	,000	-,220	-,115
	cw20m427 My job [is/was] physically demanding.	,225	,039	,162	5,738	,000	,148	,302
	nettocat Personal net monthly income in categories	-,061	,020	-,100	-3,114	,002	-,099	-,023

a. Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

**Syntax:** \*betrouwbaarheidsintervallen model 2 en 5.

REGRESSION

/DESCRIPTIVES MEAN STDDEV CORR SIG N

/MISSING LISTWISE

/STATISTICS COEFF OUTS CI(95) R ANOVA CHANGE

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

/NOORIGIN

/DEPENDENT v3d

/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat

/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat cw20m427 nettocat.

### Bijlage 3 – Assumptietoetsing en outlieranalyse

Deze bijlage geeft de controle voor multicollineariteit, assumptietoetsing en de outlieranalyse van het onderzoek weer. Indien nodig volgt er de bijbehorende output en syntax.

#### Multicollineariteit

Het is belangrijk om te weten of er sprake is van multicollineariteit in de regressieanalyse. Multicollineariteit betekent dat voorspellers te sterk met elkaar samenhangen. Zo'n hoge correlatie kan ervoor zorgen dat er problemen ontstaan, want twee onafhankelijke variabelen moeten ook echt onafhankelijk zijn. Een hulpmiddel om de multicollineariteit te meten is de Variance Inflation Factor (VIF). Over het algemeen geldt dat wanneer de VIF lager is dan 4, er geen sprake is van ernstige multicollineariteit. In de tabel hieronder is te zien dat er geen sprake lijkt te zijn van ernstige multicollineariteit. Alle variabelen hebben een VIF lager dan 4. Op één variabele na hebben alle variabelen een VIF lager dan 3.

Model		Coefficients <sup>a</sup>					Collinearity Statistics	
		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Tolerance	VIF
		B	Std. Error	Beta				
1	(Constant)	2,931	,205		14,289	,000		
	dummybouw	,223	,126	,071	1,762	,078	,440	<b>2,274</b>
	industriehandel	,067	,115	,026	,583	,560	,368	<b>2,715</b>
	semipubliek	,167	,111	,070	1,496	,135	,327	<b>3,060</b>
	geslachtdummy	-,092	,073	-,039	-1,266	,206	,758	<b>1,320</b>
	leeftijdcohort	-,036	,022	-,044	-1,631	,103	,972	<b>1,028</b>
	oplcat Level of education in CBS (Statistics Netherlands) categories	-,167	,027	-,188	-6,235	,000	,784	<b>1,276</b>
	cw20m427 My job [is/was] physically demanding.	,225	,039	,162	5,738	,000	,892	<b>1,121</b>
	nettocat Personal net monthly income in categories	-,061	,020	-,100	-3,114	,002	,688	<b>1,454</b>

a. Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

#### Syntax: \*VIF-waarden.

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT v3d
  /METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat cw20m427
  nettocat.
```

**Assumptietoetsing**

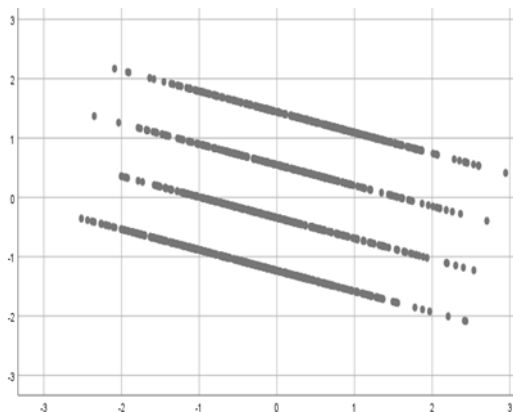
Om te kijken of het model klopt wordt er gekeken of er aan de volgende vier modelassumpties wordt voldaan.

**Assumptie 1: Onafhankelijke waarnemingen**

De eerste assumptie houdt in dat de observaties onafhankelijk van elkaar moeten zijn. Dit betekent dat de respondenten onafhankelijk van elkaar zijn getrokken in de steekproef, waardoor er geen samenhang zal ontstaan tussen de scores van de respondenten. Aan deze assumptie wordt deels voldaan. Het LISS panel maakt gebruik van geneste data, het is gebaseerd op een steekproef van huishoudens. De waarnemingen van individuele personen in het LISS panel zijn geclusterd in huishoudens, zo kunnen er meerdere individuen van hetzelfde huishouden meedoen aan het panel. Hierdoor zijn de waarnemingen niet volledig onafhankelijk van elkaar. Echter, het is een groot panel wat inhoudt dat de personen uit hetzelfde huishouden niet perse dezelfde vragenlijsten invullen. Hierdoor zijn de waarnemingen ook niet volledig afhankelijk van elkaar. Er is sprake van een gedeeltelijke schending van deze assumptie.

**Assumptie 2: Lineair verband tussen de afhankelijke variabele en de onafhankelijke variabelen**

De tweede assumptie stelt dat er een lineair verband moet zijn tussen de variabelen. Hierbij wordt gekeken naar de spreidingsdiagram die hieronder is weergegeven. Het gemiddelde voor elke set van x-waarden is niet gelijk aan 0. Er zijn duidelijk 4 categorieën in te zien. De afhankelijke variabele is een categorische variabele, waardoor een lineair model eigenlijk ongeschikt is.



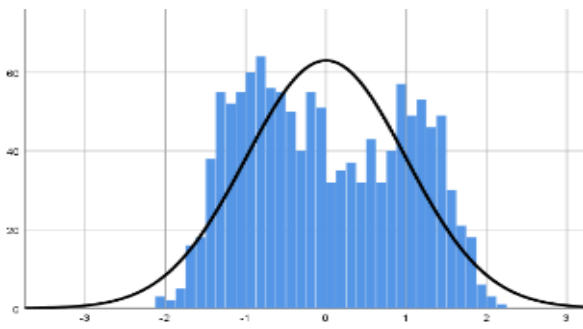
Residual plot

**Assumptie 3: Constante conditionele standaarddeviatie**

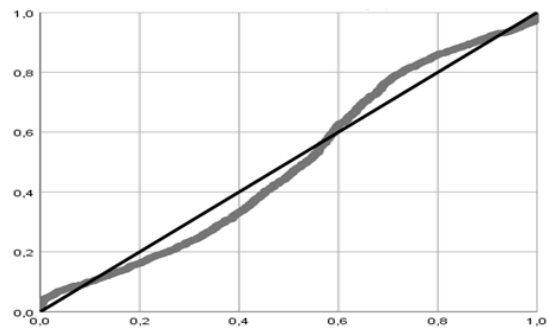
De derde assumptie stelt dat er sprake moet zijn van homoscedasticiteit. Dit betekent dat voor elke set van waarden van de x-en, y dezelfde conditionele standaarddeviatie heeft (in de populatie). Er wordt wederom gekeken naar de spreidingsdiagram. De spreiding moet bij deze assumptie overal ongeveer gelijk zijn, en dit lijkt het geval. Aan deze assumptie wordt voldaan.

**Assumptie 4: Normale verdeling van de residuen**

De laatste assumptie stelt dat de conditionele verdeling van y normaal verdeeld moet zijn. Dit wordt gecontroleerd door te kijken of de residuen normaal verdeeld zijn door middel van de histogram en PP-plot die hieronder staan weergegeven. Aan deze assumptie wordt niet voldaan. Dit komt door de afhankelijke categorische variabele en het niet passen van het lineaire model. Er zijn verschillende pieken te zien in de histogram, de verdeling is niet normaal verdeeld. Ook is er een lichte s-curve te zien in de PP-plot, wat is ontstaan door de pieken in de verdeling.



Histogram met normaalkromme



PP-plot

**Syntax:** \*spreidingsdiagram, histogram en PP-plot.

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT v3d
/METHOD=ENTER dummybouw industrieghandel semipubliek geslachtdummy leeftijdcohort oplcat cw20m427
  nettocat
/SCATTERPLOT=(*ZRESID ,*ZPRED)
/RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID)
/SAVE COOK LEVER DFBETA DFFIT.
```

**Uitbijters**

Het is ten slotte nog van belang om te kijken of er uitbijters in de data voorkomen die een eventueel vertekend beeld kunnen geven. In de histogram en de PP-plot, die op de vorige pagina zijn weergegeven, is te zien dat er sprake is van uitbijters in het onderzoek. Deze uitbijters kunnen invloed hebben op de resultaten. Zo kunnen bijvoorbeeld de hellingen worden beïnvloed door de uitbijters. Die kunnen meer positief of negatief uitvallen.

Er is door middel van de Leverage, de Cook's distance, DfBeta en DfFit gekeken welke punten de uitbijters of invloedrijke punten zijn. Er zijn ID-nummers toegevoegd, zodat er kan worden bekeken welke punten het meest opvallen.

Eerst wordt er gekeken naar invloedrijke punten in de x-richting: de Leverage. Met behulp van de formule  $3 * P/N$  is te controleren of er een probleem kan ontstaan. Hierbij is  $P(8)$  het aantal geschatte parameters en  $N(1259)$  het totale aantal respondenten. Dit geeft een Leverage-grens van: 0,01906. In de dataset wordt gekeken naar waarden die boven deze grens liggen.

De Cook's distance is een combinatie van outlier in de y-richting en de x-richting. De vuistregel hierbij is dat wanneer de Cook's distance groter is dan 1, dit een probleem is. Dit is niet het geval. Er wordt een grenswaarde uitgerekend ( $4/N = 0,0032$ ). Er wordt gekeken naar waarden die boven deze grens liggen.

Met de DfBeta kijkt men hoe ver de coëfficiënt verandert als de regressieanalyse opnieuw wordt gedaan zonder de betreffende case. De DfBeta wordt voor elke parameter uitgerekend. Er is gekeken naar de grootste positieve en negatieve DfBeta's die voor de grootste invloed kunnen zorgen.

Met de DfFit wordt gekeken hoe ver y verandert als de regressieanalyse opnieuw wordt uitgevoerd zonder de betreffende case. De grootste negatieve DfFit waarde is -0,02950. De grootste positieve DfFit waarde is 0,04891. Het zijn erg kleine waarden, wanneer de betreffende cases worden verwijderd uit de dataset zal de helling dan ook niet erg worden beïnvloed.

In de tabel hieronder zijn de verschillende maten te zien met betreffende scores en ID-nummers. De individuen die op meerdere maten opvallen zijn de individuen met de nummers: 654, 875, 142, 331, 385, 774 en 776. Deze individuen zullen uit de dataset worden gehaald. De regressieanalyse wordt opnieuw uitgevoerd om te kijken of het iets veranderd aan de hellingen en significantie.

**Syntax: \*controlerenoutliers&invloedrijkepunten.**

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT v3d
  /METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat cw20m427
  nettocat
  /SCATTERPLOT=(*ZRESID , *ZPRED)
  /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID)
  /SAVE COOK LEVER DFBETA DFFIT.
```

Leverage	ID	Cook's	ID	DfBeta	ID	DfFit	ID
0,02538	654	0,00813	875	-0,02186	875	0,04891	654
0,02291	1169	0,00808	654	-0,01792	1213	0,03992	331
0,02170	871	0,00740	331	-0,01749	474	0,03960	142
0,02001	142	0,00669	385	-0,01637	1215	0,03723	875
0,01887	1033	0,00666	142	0,01598	774	0,03595	385
0,01824	331	0,00617	776	-0,01585	654	-0,02950	776
0,01818	489	0,00543	237	0,01583	304	-0,02572	774
0,01800	55	0,00530	1247	-0,01555	142	0,02557	65

Na het verwijderen van de cases is in de tabel hieronder te zien dat er geen opmerkelijke veranderingen zijn in de hellingen of de significantie. De uitbijters lijken geen invloed te hebben op de uitkomsten van de regressieanalyse.



**Syntax: \*casenummersaanmaken.**

```
COMPUTE id=$CASENUM.
FORMAT id (F8.0).
EXECUTE.
```

**Syntax: \*verwijderencases.**

```
USE ALL.
COMPUTE filter_$=((id ~= 654 & id ~= 875 & id ~= 142 & id ~= 331 & id ~= 385 & id ~= 774 & id ~= 776)).
VARIABLE LABELS filter_$ '(id ~= 654 & id ~= 875 & id ~= 142 id ~= 331 & id ~= 385 & id ~= 774 & id ~= 776 (FILTER))'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized	t	Sig.
		B	Std. Error	Coefficients Beta		
1	(Constant)	2,942	,205		14,376	,000
	oplcat Level of education in CBS (Statistics Netherlands) categories	-,177	,027	-,200	-6,616	,000
	nettocat Personal net monthly income in categories	-,064	,020	-,104	-3,224	,001
	dummybouw	,268	,127	,086	2,109	,035
	industriehandel	,102	,116	,039	,885	,376
	semipubliek	,207	,112	,087	1,843	,065
	geslachtdummy	-,074	,072	-,031	-1,017	,309
	leeftijdcohort	-,038	,022	-,047	-1,732	,083
	cw20m427 My job [is/was] physically demanding.	,227	,039	,163	5,769	,000

a. Dependent Variable: v3d Als u denkt aan werken tot uw AOW-leeftijd, in hoeverre bent u boos dat u zo lan

**Syntax: \*regressiezonderdeuitbijters.**

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT v3d
/METHOD=ENTER dummybouw industriehandel semipubliek geslachtdummy leeftijdcohort oplcat cw20m427
nettocat
/SCATTERPLOT=(*ZRESID ,*ZPRED)
/RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID)
/SAVE COOK LEVER DFBETA DFFIT.
```