

POLITICAL PARTICIPATION IN THE LIGHT OF THE GOALFRAMING THEORY

Politieke participatie in het licht van de goalframe theorie

Welke goalframe is dominant in de keuze van individuen om politiek te gaan participeren?

1-8-2023

Jelte Gerrit de Vries

S2574365

j.g.de.vries.4@student.rug.nl

Master scriptie

Master Sociologie van arbeidsrelaties en levensloop

Rijksuniversiteit Groningen

Scriptiebegeleider: Menno Rol

Referent: Andreas Flache



Abstract

Dit onderzoek probeert de goalframes van Lindenberg en Steg (2007) te gebruiken om verschillen tussen zelfstandigen en werknemers in politieke participatie te verklaren. De *gain goalframe* zou ervoor zorgen dat mensen met grotere persoonlijke belangen eerder gaan participeren, terwijl de *normative goalframe* ervoor zorgt dat mensen gaan politiek participeren wanneer dat als sociaal wenselijk wordt ervaren. In dit onderzoek proberen we verschillen in politieke participatie aan te tonen die ontstaan door verschillen in dominante goalframes tussen werknemers en zelfstandigen. We beredeneren dat zelfstandigen gevoeliger zijn voor de *gain goalframe*, omdat zij grotere persoonlijke belangen hebben. Daarnaast verwachten we met de *normative goalframe* dat de invloed van opleidingsniveau op politieke participatie groter is voor werknemers dan voor zelfstandigen. De redenering is dat hoogopgeleide werknemers meer dan hoogopgeleide zelfstandigen geprikkeld worden door hun omgeving waarin meer andere hoogopgeleiden bevinden die van zichzelf al een meer positieve houding ten opzichte van politieke participatie hebben. Tot slot onderzoeken we of die invloeden verschillend zijn voor actieve dan voor passieve vormen van politieke participatie.

Om dat te onderzoeken zijn er binaire logistische modellen geschat met data van het LISS-panel uit 2022. Deze data bestaat uit een representatieve steekproeftrekking van de Nederlandse bevolking. In de modellen zijn actieve en passieve politieke participatie de afhankelijke variabelen met opleidingsniveau en werksituatie als verklarende variabelen. Daarnaast gebruiken we migratieachtergrond, leeftijd, geslacht en sector als controlevariabelen. De resultaten van het onderzoek tonen geen verschillen tussen werknemers en werkgevers, ook niet op basis van opleidingsniveau. Dit interpreteren we als een indicatie dat er geen verschillen zijn in welke goalframe dominant is tussen zelfstandigen en werknemers. Wel zien we dat er verschillen zijn in de invloeden van andere variabelen op actieve en passieve vormen van politieke participatie. De aanbeveling is daarom om de verschillen tussen actieve en passieve politieke participatie verder te onderzoeken, bijvoorbeeld door te kijken of er verschillen zijn in de mensen die doen aan passieve politieke participatie en actieve politieke participatie. Zijn dat dezelfde groepen, of zijn dit juist verschillende groepen? Ander onderzoek zou met een kwalitatieve onderzoeksmethode gericht kunnen vragen naar de beweegredenen voor politieke participatie, zodat de goalframes meer direct in beeld kunnen worden gebracht.

Inhoud

1. Inleiding	4
2. Theoretisch kader	6
2.1 Goalframes	6
2.2 Gain goalframe	7
2.3 Normative goalframe.....	9
2.4 Opleidingsniveau.....	10
2.5 Controle variabelen	12
3. Methode.....	13
3.1 Data en steekproef.....	13
3.2 Operationalisatie van de variabelen.....	13
3.3 Analyseopzet	16
4. Resultaten	19
4.1 Beschrijvende statistieken	19
Univariate statistieken	19
Bivariate statistieken	20
4.2 Modevaluatie	22
Assumptie	22
Outliers en invloedrijke punten	23
Multicollineariteit.....	24
Fit van de modellen	24
4.3 Hypothesetoetsing	27
5 Conclusie & discussie	36
5.1 Conclusie	36
5.2 Discussie.....	38
6. Bronvermelding.....	41
Bijlage 1 Operationalisatie van de variabelen	44
Bijlage 2: Analyses	60
Univariate statistieken	60
Bivariate statistieken	63
Multivariate statistieken	77
Bijlage 3 Modevaluatie.....	86
Assumpties	86

Outliers en invloedrijke punten	86
Multicollineariteit.....	101
Modelfit.....	103

1. Inleiding

Politieke participatie is nodig zodat een democratie goed kan functioneren (Van der Meer & Dekker, 2011). Het zorgt er namelijk voor dat de belangen van burgers vertegenwoordigd worden. Maar, een democratie functioneert alleen goed als de politieke participatie gelijk gespreid is over de populatie, zodat de belangen van iedereen in gelijke mate worden vertegenwoordigd. Anders komen de belangen van de een sterker naar voren dan die van de ander. Het is daarom van belang om te onderzoeken op welke manier we politieke participatie kunnen verklaren, zodat we beleid gericht kunnen inzetten om de politieke participatie van achtergestelde groepen te verhogen.

Om politieke participatie te verklaren gebruiken we nu vaak economische theorieën of de mobilisatietheorie (Leighley, 1995). Deze economische theorieën stellen dat mensen politiek participeren wanneer ze daarmee hun persoonlijke belangen kunnen behartigen. De mobilisatietheorie stelt daarnaast dat mensen gaan participeren wanneer ze voldoende door hun omgeving worden geprikkeld. Deze theorieën kunnen we spiegelen aan de *goalframing theory* (Lindenberg & Steg, 2007) die in de sociologie vaak wordt gebruikt om het gedrag van mensen te verklaren. Deze theorie stelt dat mensen verschillende goalframes hebben die op verschillende momenten dominant kunnen zijn. De goalframe die dominant is, bepaalt welk gedrag mensen vertonen. Deze goalframes zorgen namelijk voor verschillen in doelen en de manieren waarop mensen worden geprikkeld. Die goalframes zijn: *hedonic*, *gain* en *normative*. De *hedonic goalframe* leidt ertoe dat mensen op zoek gaan naar een goed gevoel op de korte termijn. De *gain goalframe* zorgt ervoor dat mensen hun middelen willen behouden of vergroten op de lange termijn. Tot slot bepaalt de *normative goalframe* dat mensen zich gaan gedragen naar wat sociaal wenselijk wordt bevonden. Wanneer we kunnen achterhalen welke goalframe in welke situatie dominant is, kunnen we dus meer zeggen over welke factoren van invloed zijn op het gedrag van mensen.

Deze goalframes kunnen we spiegelen aan de eerder genoemde theorieën: als de *gain goalframe* dominant is, dan zou dat betekenen dat mensen gaan politiek participeren, wanneer ze daarmee hun middelen kunnen vergroten of behouden. Daarmee past de *gain goalframe* goed bij de economische theorieën. De *normative goalframe* zou ervoor zorgen dat mensen politiek participeren wanneer andere mensen van ze verwachten dat ze gaan participeren, dus wanneer mensen door hun sociale context worden geprikkeld om te gaan participeren. De *normative goalframe* past daardoor bij de mobilisatietheorie.

In dit onderzoek kijken we of er een indicatie is dat er verschillen zijn tussen groepen in de manier waarop de goalframes een rol spelen bij politieke participatie. De groepen die we in dit onderzoek vergelijken zijn werknemers en zelfstandigen. Met zelfstandigen bedoelen we mensen met een eenmanszaak: ondernemers die in hun eentje in het bezit zijn van een onderneming met of zonder personeel. We verwachten dat bij zelfstandigen de *gain goalframe* relatief dominanter is, terwijl bij de werknemers de *normative goalframe* relatief dominanter is. Wanneer we aan kunnen tonen dat de ene goalframe een grotere invloed heeft op de politieke participatie dan de andere, kunnen we meer inzicht verkrijgen in de manier waarop we politieke participatie kunnen verklaren. De onderzoeksvraag die we daarom met dit onderzoek proberen te beantwoorden is:

In hoeverre kunnen we politieke participatie verklaren door verschillen in ondernemerschap of werknemerschap en opleidingsniveau?

Hoewel er veel variatie is in vormen van politieke participatie, wordt het vaak als één concept gedefinieerd (Leighley, 1995). Het is aannemelijk dat verbanden verschillend werken voor verschillende vormen van politieke participatie. Zo vraagt het meelopen aan een demonstratie veel meer tijd en inzet dan het stemmen voor de Tweede Kamerverkiezingen. In dit onderzoek onderscheiden we daarom actieve en passieve politieke participatie. Met actieve politieke participatie bedoelen we vormen van politieke participatie die meer inzet vragen van het individu en die ook meer zichtbaar zijn. Voorbeelden daarvan zijn demonstreren en het spreken op een bijeenkomst. Passieve politieke participatie staat voor vormen van politieke participatie waarbij er weinig inzet wordt gevraagd en de participatie minder zichtbaar is. Door deze twee vormen van politieke participatie met elkaar te vergelijken kunnen we bestaande theorieën over politieke participatie aanvullen. Een subvraag die we in dit onderzoek hanteren is daarom:

Wat zijn de verschillen tussen actieve en passieve vormen van politieke participatie?

In dit onderzoek onderzoeken we de hoofdvraag en subvraag aan de hand van binaire logistische regressiemodellen. Hiervoor toetsen we de hypothesen die voortkomen uit de theorie. De data die we gebruiken is afkomstig van het LISS-panel uit 2022. We kijken daarbij naar interactie-effecten tussen zelfstandigen met hoog opleidingsniveau op zowel actieve als politieke participatie. Er is gecontroleerd voor leeftijd, geslacht, migratieachtergrond en of iemand in de overheidssector werkt.

2. Theoretisch kader

Demonstreren, stemmen en het schrijven van een brief aan de gemeenteraad zijn allemaal duidelijke vormen van politieke participatie. Desondanks is de definitie van politieke participatie niet zo helder als die in het eerste opzicht lijkt. Er is namelijk discussie over of bijvoorbeeld vrijwilligerswerk of een politieke discussie voeren met vrienden en familie ook onder politieke participatie vallen (Sairambay, 2020). Een veelgebruikte definitie is die van Brady, Verba en Schlozman (1995): “een activiteit die bedoeld is om of als gevolg heeft, indirect of direct overheidsoptreden te beïnvloeden”. Het nadeel van deze definitie wordt geformuleerd door Sairambay (2020) en dat is dat elke activiteit die een invloed heeft op overheidsoptreden met deze definitie gezien kan worden als politieke participatie, terwijl daarmee ook activiteiten worden inbegrepen die niet noodzakelijk bedoeld zijn om overheidsoptreden te beïnvloeden. Volgens deze definitie hoeft een gedraging dus niet noodzakelijk het doel te hebben om het overheidsoptreden te beïnvloeden, om aangemerkt te worden als politieke participatie. Dat is problematisch, omdat het gedrag van burgers heel vaak overheidsoptreden beïnvloedt, ook al is dat niet de bedoeling. Criminaliteit is bijvoorbeeld ook gedrag van mensen dat overheidsoptreden kan beïnvloeden, maar dat kunnen we moeilijk beschouwen als politieke participatie. Daarom introduceren we in dit onderzoek de volgende definitie van politieke participatie: het door burgers ondernemen van of meedoen aan activiteiten die als doel hebben om politieke besluiten of processen te beïnvloeden. Binnen deze kaders kunnen we vervolgens onderscheid maken tussen verschillende vormen van participatie. Enerzijds passieve participatie, waarbij een activiteit weinig inzet vraagt en daardoor ook minder zichtbaar is. Hieronder kunnen we onder andere het uitbrengen van een stem of het tekenen van een petitie scharen. Anderzijds kunnen we actieve vormen van participatie onderscheiden die meer inzet vragen en meer zichtbaar zijn, zoals demonstreren, discussiëren in een panel of op een bijeenkomst, of contact opnemen met politici.

2.1 Goalframes

Om politieke participatie te verklaren kunnen we de goalframes van Lindenberg en Steg (2007) hanteren. We verwachten namelijk dat er tussen groepen respondenten verschillende goalframes dominant zijn. De goalframes beschrijven welke doelen bepalen hoe individuen zich gaan gedragen en welke prikkels daarbij een rol spelen. In de theorie van Lindenberg en Steg zijn drie verschillende goalframes onderscheiden: *hedonic goalframe*, *gain goalframe* en

normative goalframe. Met de *hedonic goalframe* zijn individuen gericht op het bereiken van een goed gevoel op de korte termijn. Met de *gain goalframe* probeert het individu zijn of haar middelen efficiënt in te zetten om die middelen te verbeteren, te behouden of nieuwe middelen te verkrijgen. Ten slotte zorgt de *normative goalframe* ervoor dat mensen erop gericht zijn om zich te gedragen naar wat sociaal wenselijk wordt bevonden. Welke goalframe in een situatie dominant is, bepaalt dus voor welke prikkels individuen gevoelig zijn en welke overweging er gemaakt wordt. Met de *gain goalframe* en met de *normative goalframe* kunnen we ook verklaringen omschrijven voor politieke participatie.

2.2 Gain goalframe

Allereerst richten we ons op de *gain goalframe*. Zoals eerder genoemd, probeert het individu met de *gain goalframe* diens economische belangen te behartigen en haar/zijn middelen efficiënt in te zetten om die middelen te verbeteren, te behouden of nieuwe middelen te verkrijgen. Kort gezegd maken individuen een economische overweging. Allereerst betekent dit dat wanneer de individuele belangen hoger zijn, mensen eerder zullen gaan participeren. De gepercipieerde opbrengsten zijn dan namelijk hoger. Daarnaast betekent dit dat mensen alleen zullen participeren wanneer zij denken dat hun bijdrage een verschil kan maken in de uiteindelijke uitkomst, want anders is de inzet van middelen niet efficiënt. Dit sluit aan bij economische theorieën die politieke participatie proberen te verklaren, zoals omschreven door Leighley (1995).

Vanuit deze invalshoek gekeken naar de politieke participatie van werknemers en zelfstandigen, dan zouden we mogen verwachten dat zelfstandigen eerder politiek participeren dan werknemers. Zij hebben namelijk grotere gepercipieerde belangen dan werknemers. Dat komt doordat politieke beslissingen meer zichtbare invloed hebben op de economische positie van zelfstandigen dan van werknemers. Politieke beslissingen over bijvoorbeeld arbeidsongeschiktheidsverzekeringen, pensioenen en Arbo-regels zijn namelijk van directe invloed op de arbeidsomstandigheden en de kosten van de zelfstandige, terwijl voor een werknemer al veel meer van deze bepalingen en voorzieningen zijn vastgelegd in cao's die niet door politieke besluiten zijn bepaald, maar door de sociale partners. Weliswaar zijn er ook belangenorganisaties die opkomen voor de belangen van zelfstandigen, maar de genomen besluiten over zelfstandigen blijven politieke besluiten, terwijl onderhandelingen voor sociale akkoorden zich grotendeels buiten de politiek afspelen. Terwijl politieke beslissingen meer zichtbare invloed hebben op de arbeidsomstandigheden van zelfstandigen, zijn eventuele veranderende kosten als gevolg van de politieke besluiten vaak niet zichtbaar voor de

werknemer, terwijl de zelfstandige ze zelf moet betalen. Eventuele politieke uitkomsten zijn daardoor meer zichtbaar en van grotere invloed op de economische positie voor de zelfstandige dan voor de werknemer.

Naast hun eigen economische belangen hebben zelfstandigen meer baat bij het economische welzijn van hun omgeving (Rotolo & Wilson, 2006). Zelfstandigen zijn in economisch opzicht namelijk vaker afhankelijk van hun omgeving, omdat deze omgeving hun klanten, leveranciers en investeerders kan bevatten. Een economisch welvarende omgeving kan dus zorgen voor een betere markt voor de zelfstandige, terwijl de werknemer in economisch opzicht vooral afhankelijk is van de werkgever. Daardoor zijn zelfstandigen meer gemotiveerd om op te komen voor de belangen van hun omgeving dan werknemers (Mencken, Smith & Tolbert, 2020).

Naast hogere gepercipieerde baten is het ook van belang dat mensen geloven dat ze met politieke participatie hun belangen kunnen behartigen, dus dat hun inzet efficiënt is. Wanneer mensen namelijk verwachten dat hun bijdrage geen of weinig verschil zal maken, dan is de beoogde opbrengst van het gedrag lager en wegen die niet op tegen de verwachte kosten. Het geloof van het individu dat hij of zij in staat is om invloed uit te oefenen op politieke besluiten of processen wordt politieke *self-efficacy* genoemd (Becker, 2004; Hadjar & Beck, 2010). Deze *efficacy* valt op te delen in twee soorten: interne *efficacy* is de mate waarin het individu denkt dat hij of zij de middelen en de vaardigheden heeft om politieke invloed uit te oefenen (Hadjar & Beck, 2004) en externe *efficacy* staat voor de mate waarin het individu erop vertrouwt dat het politieke proces ontvankelijk is voor de input van het individu (Becker, 2004). Zelfstandigen hebben een hogere interne *efficacy* dan werknemers, doordat ze minder gestandaardiseerd werk hebben en daardoor meer vaardigheden ontwikkelen die relevant zijn voor politieke participatie (Wilson & Musick, 1997). Met die verbeterde vaardigheden zijn zelfstandigen in verhouding tot werknemers dus beter in staat om invloed uit te oefenen, waardoor ze ook eerder geloven dat ze meer invloed kunnen uitoefenen. De gepercipieerde opbrengst is voor een zelfstandige dus hoger dan voor een werknemer, waardoor een zelfstandige eerder gemotiveerd zal zijn om te participeren.

De rol van de overweging tussen kosten en baten zal waarschijnlijk meer aanwezig zijn bij actieve vormen van politieke participatie dan voor passieve vormen. Dat komt doordat passieve vormen in verhouding tot actieve vormen weinig inzet en kosten vragen, waardoor de opbrengsten eerder opwegen tegen de kosten. Omdat het weinig inzet vraagt, spelen het

hebben van vaardigheden en kennis ook minder een rol bij het participeren. Dus zal de interne *efficacy* een minder grote rol spelen bij passieve vormen, dan bij actieve vormen.

Toch heeft deze economische benadering een belangrijke beperking. Namelijk, in de meeste gevallen is de marginale opbrengst van politieke participatie van een individu verwaarloosbaar. Doordat er zoveel mensen betrokken zijn in het democratische proces is de kans dat de politieke participatie van één individu het verschil maakt zeer klein. Dat is omdat de kans dat de stem van één iemand invloed gaat hebben op de uitslag van de verkiezingen zeer klein is, maar ook omdat de kans dat een extra individu meeloopt in een demonstratie die een politiek besluit gaat beïnvloeden erg klein is. Weliswaar kan iemand redelijk grote invloed uitoefenen door een demonstratie te organiseren of een burgerinitiatief te starten. De marginale opbrengst van dit individu zou dan nog relatief groot zijn, maar deze demonstraties en burgerinitiatieven zouden alleen succesvol zijn als er meer individuen bij aansluiten. Voor die individuen aansluiten bij die initiatieven is de marginale opbrengst dan al weer veel kleiner. Toch zien we dat mensen veel participeren, dus de politieke participatie is hoger dan we zouden verwachten op basis van individuele kosten-baten analyses. Er lijkt dus een andere goalframe een rol te spelen: de *normative goalframe*.

2.3 Normative goalframe

Er zijn dus meer prikkels die een rol spelen dan alleen de economische. Hiervoor kunnen we een ander goalframe van Lindenberg & Steg gebruiken (2007): het *normative goalframe*. Binnen dit goalframe gedragen mensen zich naar wat zij denken dat passend is en wat door andere mensen als positief zal worden beschouwd. Mensen krijgen dus prikkels uit hun sociale omgeving die aangeven welk gedrag wenselijk is, wat vervolgens de gedraging bepaalt. Dit sluit aan bij de mobilisatietheorie. Deze theorie beschrijft dat mensen politiek gaan participeren, wanneer ze daarvoor voldoende prikkels krijgen (Leighley, 1995). Deze prikkels kunnen heel verschillend van aard zijn: dit kunnen bijvoorbeeld de verkiezingscampagne van een politieke partij zijn, een nieuwsbericht over een politieke kwestie of de verkiezingen (Rosenstone & Hansen, 1993) of een informele politieke discussie met vrienden, familie of collega's (Gilbert, 1993; Kenny, 1992; Leighley, 1990). Met deze prikkels wordt het voor mensen zichtbaar gemaakt welk gedrag als positief wordt beschouwd, waardoor ze gemotiveerd worden om dat gedrag te vertonen. Wanneer zichtbaar is welk gedrag als positief wordt beschouwd, ontstaan er verwachtingen. Wanneer er in de omgeving van het individu van mensen verwacht wordt dat ze participeren is de kans op participatie groter dan wanneer de omgeving die verwachtingen niet heeft (Ostrom, 2000). Wanneer

mensen niet aan de verwachtingen voldoen, kunnen ze namelijk te maken krijgen met sociale sancties.

Mensen willen namelijk graag voldoen aan die verwachtingen om sociale sancties te ontwijken. Wanneer de verwachtingen er niet zijn, zullen waarschijnlijk de sociale sancties ook lager zijn. Tegelijkertijd kan het voldoen aan verwachtingen beschouwd worden als pro-sociaal gedrag, waardoor er met het vertonen van dat gedrag een betere reputatie kan worden verkregen (Milinski, 2016).

De invloed van deze sociale prikkels is waarschijnlijk sterker voor actieve vormen dan voor passieve vormen van politieke participatie, omdat actieve vormen zichtbaarder zijn. Zichtbaarheid zorgt er namelijk voor dat het gedrag van mensen kan worden beoordeeld door andere mensen en dus onder invloed staan van sociale mechanismes die het pro-sociale gedrag bevorderen (Ostrom, 2000).

2.4 Opleidingsniveau

Van hoogopgeleiden weten we al dat ze meer politiek participeren dan laagopgeleiden (Milligan, Moretti & Oreopoulos, 2004). Ze zijn namelijk meer gesocialiseerd met politieke waarden die politieke participatie aanmoedigen (Lind, 2006; Nie, Junn & Stehlik-Barry, 1996) en ze beschikken over meer middelen zoals kennis en vaardigheden die ze kunnen inzetten (Dalton, 1984). Deze middelen zorgen voor een hogere interne politieke *self-efficacy* (Hadjar & Beck, 2010) en ze hebben een hoger sociaal kapitaal dat ervoor zorgt dat ze ook meer middelen vanuit hun netwerk kunnen aanwenden om te participeren (Putnam, 2000). Maar daarnaast zullen hogeropgeleiden ook meer prikkels in hun omgeving tegenkomen die ze mobiliseren om te participeren. Dat komt doordat hogeropgeleiden zich vaker in een omgeving bevinden waarin politieke participatie voorkomt (Leighley, 1995).

Hogeropgeleiden gaan namelijk vaker om met andere hoogopgeleiden, bijvoorbeeld doordat ze vrienden zijn geworden met studie- of klasgenoten of doordat ze vaak werken met collega's met min of meer dezelfde opleiding. Hoogopgeleiden hebben daarom vaker interacties met andere hoogopgeleiden, waardoor ze meer prikkels krijgen vanuit hun omgeving die aansturen op politieke participatie. Voor hoog opgeleiden zorgt de *normative goalframe* er dus voor dat gedrag wordt bijgestuurd naar het sociaal wenselijke gedrag van politieke participatie.

Bij zelfstandigen kunnen we een ander verband verwachten tussen de opleidingsniveau en politieke participatie dan bij werknemers. We verwachten namelijk dat het mobiliserende

effect van opleidingsniveau groter is voor werknemers dan voor zelfstandigen. Werknemers werken namelijk vaak samen met collega's en leidinggevenden die min of meer dezelfde opleiding hebben gevolgd en dus eerder hetzelfde denken over politieke participatie. Deze relaties ontbreken vaak bij zelfstandigen. Zelfstandigen kunnen weliswaar werknemers hebben, maar door de verticale relatie die zij met elkaar hebben is de kans dat de werknemer de zelfstandige beïnvloedt kleiner dan dat werknemers elkaar onderling beïnvloeden. Werknemers hebben dus naar verwachting meer interacties met mensen met dezelfde opleiding dan zelfstandigen, waardoor voor hen de invloed van opleidingsniveau groter is dan voor zelfstandigen. Dat betekent voor dat hoogopgeleide werknemers eerder door hun werkomgeving zullen worden geprikkeld om politiek te participeren dan hoogopgeleide zelfstandigen. Hierdoor zijn hoogopgeleide werknemers meer gemotiveerd om politiek te participeren dan hoogopgeleide zelfstandigen. Voor laag opgeleiden zullen deze verschillen kleiner zijn, doordat in de werkomgeving van laagopgeleiden waarschijnlijk het vertonen van politieke participatie minder wordt gepromoot en daardoor ook minder wordt besproken. Het verschil in prikkels tussen laagopgeleide zelfstandigen en werknemers is daardoor kleiner.

Hoogopgeleide werknemers zullen daardoor eerder gemobiliseerd worden dan hoogopgeleide zelfstandigen. Tegelijkertijd zullen de verschillen tussen laagopgeleide werknemers en zelfstandigen minder groot zijn, doordat de werkomgeving van laagopgeleiden in een lagere mate politieke participatie mobiliseert. Doordat mobilisatie waarschijnlijk een belangrijkere factor is voor actieve vormen dan voor passieve vormen van politieke participatie, zullen de verschillen daarin groter zijn bij werknemers dan bij zelfstandigen.

Uit deze theorie volgen de volgende hypothesen:

H1) Het effect van opleidingsniveau is sterker voor werknemers dan voor zelfstandigen op passieve politieke participatie.

H2) Het effect van opleidingsniveau is sterker voor werknemers dan voor zelfstandigen op actieve politieke participatie.

H3) Hoog opgeleide werknemers hebben een hogere mate van passieve politieke participatie dan hoog opgeleide zelfstandigen.

H4) Hoog opgeleide werknemers hebben een hogere mate van actieve politieke participatie dan hoog opgeleide zelfstandigen.

H5) De invloed van opleidingsniveau en zelfstandig ondernemerschap is groter voor actieve vormen van politieke participatie dan voor passieve vormen.

2.5 Controle variabelen

Om deze hypothesen te kunnen toetsen is het van belang om de oorzaken te kunnen isoleren van alternatieve oorzaken die een rol kunnen spelen. Wanneer die alternatieve oorzaken van politieke participatie verschillend verdeeld zijn over zelfstandigen en werknemers of over opleidingsniveaus kunnen ze namelijk de uitkomsten van het onderzoek beïnvloeden.

Allereerst kan een verschil in leeftijd voor verschillen in participatie zorgen: volgens Oppenhuis (1995) gaan mensen met een hogere leeftijd meer participeren doordat met de leeftijd politieke betrokkenheid, interesse en kennis stijgen. Dit terwijl zelfstandigen gemiddeld ouder zijn dan werknemers (Sociaal Economische Raad, 2010). Maar ook geslacht kan voor verschillen in politieke participatie zorgen, al zijn die verschillen per land anders in grootte en richting (Oppenhuis, 1995). Bij werknemers zijn de geslachten bijna gelijk verdeeld, maar bij zelfstandigen zijn mannen in de meerderheid, 62% tegenover 38% vrouwen (Centraal Bureau voor de Statistiek, 2021a). Daarnaast werken zeer weinig zelfstandigen bij de overheid of semi-overheden (CBS, 2021b), terwijl mensen binnen de publieke sector eerder politiek participeren dan mensen die in de private sector werken (Brewer, 2003; Ertas, 2015). Tot slot wordt voor migratieachtergrond gecontroleerd, omdat er verschillen zijn tussen de mate van participatie op basis van deze migratieachtergrond (CBS, 2021): eerste generatie immigranten doen het minst aan politieke acties (38,5%), terwijl de tweede generatie juist meer meedoet (48,5%) aan politieke acties dan mensen zonder migratieachtergrond (45,7%). Bovendien zijn mensen met een migratieachtergrond met 17% van de werkzame beroepsbevolking wat minder snel een zelfstandige dan mensen zonder migratieachtergrond met 19% (CBS, 2022).

3. Methode

In dit hoofdstuk worden de onderzoeksmethoden besproken. Daarvoor beschrijven we in paragraaf 3.1 eerst de gebruikte dataset en de steekproef die daarop is uitgevoerd. Vervolgens leggen we in paragraaf 3.2 uit welke variabelen zijn gebruikt tot slot bespreken we in 3.3 op welke manier de hypothesen zullen worden getoetst.

3.1 Data en steekproef

Voor dit onderzoek is data gebruikt van het LISS-panel van CentERdata. Dit panel bestaat uit data van Nederlandse respondenten die komen uit een representatieve steekproef van huishoudens uit het bevolkingsregister (Centerdata, z.d.): Deze aselechte steekproef bevat zo'n 7.500 individuen vanaf 16 jaar uit ongeveer 5.000 huishoudens. De respondenten worden gevraagd om maandelijks een online vragenlijst in te vullen waar ze vervolgens een geldelijke bijdrage voor krijgen. De respons hangt voor elke vragenlijst op zo'n 80% van de geregistreerde huishoudens. Er zijn acht thema's die met de vragenlijsten onderzocht worden: gezondheid, politiek en waarden, geloof en etniciteit, vrijetijdsbesteding, familie en huishouden, werk en scholing, persoonlijkheid en tenslotte economische situatie. Het beantwoorden vragenlijsten kost tussen de 15 en 30 minuten tijd. Respondenten die anders niet zouden kunnen participeren, worden voor het onderzoek van een computer met internetverbinding voorzien.

Voor dit onderzoek gebruiken we de vragenlijsten van 'politiek en waarden' en 'werk en scholing' gecombineerd met de achtergrondkenmerken van de respondenten (wave 14, april '22). Na het samenvoegen van de databestanden en het selecteren op respondenten die waarden hebben op alle gebruikte variabelen én respondenten van 18 jaar of ouder blijft er een steekproef over van 2604, waarvan 199 zelfstandigen. De steekproef blijft in verdeling over de variabelen redelijk gelijk, alleen leeftijd en opleidingsniveau veranderen. Dat valt te verklaren doordat alleen werkenden in de steekproef komen, waardoor jonge mensen en gepensioneerden erbuiten vallen. Daardoor vallen ook veel mensen zonder afgeronde opleiding, geclassificeerd als laagopgeleid, buiten onze analyse.

3.2 Operationalisatie van de variabelen

In deze paragraaf zetten we uiteen hoe de variabelen zijn geoperationaliseerd. We benoemen voor elke variabele de originele vraagstelling, de antwoordopties, de bewerking en de uiteindelijke variabele.

Passieve politieke participatie

Deze variabele bestaat uit één vraag, namelijk: *Hebt u tijdens de laatste Tweede Kamerverkiezingen op 17 maart 2021 gestemd?* Er zijn geen andere variabelen in de dataset die passieve politieke participatie kunnen vertegenwoordigen. Bij de vraag konden de respondenten antwoorden ‘ja’, ‘nee’ of ‘niet-stemgerechtigd’. In de uiteindelijke variabele zijn de categorieën: 1) passief geparticipeerd en 0) niet passief geparticipeerd.

Actieve politieke participatie

Deze variabele bestaat uit zeven verschillende items. Voor deze items werd gevraagd: *Wilt u van de volgende mogelijkheden aangeven of u deze in de afgelopen 5 jaar hebt gebruikt? (Meerdere antwoorden mogelijk)*. De antwoordcategorieën zijn voor elk item: ‘nee’, ‘ja’ of ‘weet ik niet’.

1. *Hulp gevraagd aan radio, televisie of krant*
2. *Politieke partij of organisatie ingeschakeld*
3. *Meegedaan aan een door de overheid georganiseerde inspraakbijeenkomst, hoorzitting of discussiebijeenkomst*
4. *Contact opgenomen met een politicus of ambtenaar*
5. *Meegedaan aan een actiegroep*
6. *Meegedaan aan een protestactie, protestmars of demonstratie*
7. *Via internet, e-mail of sms meegedaan aan een politieke discussie of actie.*

In de uiteindelijke variabele heeft elke respondent een waarde van 1) meegedaan aan tenminste één van de mogelijkheden van actieve politieke participatie of 0) niet meegedaan aan elke mogelijkheid van actieve politieke participatie.

Opleidingsniveau

Voor ‘opleidingsniveau’ is een ingelezen achtergrondvariabele gebruikt die aangeeft wat de hoogst genoten opleiding met diploma van de respondent is. De categorieën waren: ‘basisonderwijs’, ‘vmbo’, ‘havo/vwo’, ‘mbo’, ‘hbo’, ‘wo’, ‘anders’, ‘(nog) geen onderwijs afgerond’ en ‘volgt nog geen onderwijs’. Een aantal categorieën waren te klein om mee te nemen in de analyse, dus zijn van de variabele twee dummyvariabelen gemaakt: *laagopgeleid* en *hoogopgeleid*. Gemiddeld opgeleid is dus de referentiecategorie. Onder laagopgeleid vallen de oorspronkelijke categorieën ‘(nog) geen onderwijs afgerond’, ‘volgt nog geen onderwijs’, ‘basisonderwijs’. Hoogopgeleid omvat de categorieën ‘hbo’ en ‘wo’. Tot slot

omvat de referentiegroep, die we omschrijven als gemiddeld opleidingsniveau, de categorieën ‘havo/vwo’ en ‘mbo’.

Zelfstandige

Deze variabele is opgebouwd uit twee variabelen. Bij de eerste variabele is gevraagd: *Bent u werknemer in vaste of tijdelijke dienst, oproepkracht, uitzendkracht of bent u een zelfstandige/freelancer of vrije beroepsbeoefenaar? Het gaat hier om de baan waarin u de meeste uren werkt, (andere) (bij-)banen komen later aan de orde. Had u meerdere (hoofd)banen die wat uren betreft gelijk zijn, neem dan de baan die voor u het belangrijkste is.* De antwoordopties waren: werknemer in vaste dienst, werknemer in tijdelijke dienst, oproepkracht, uitzendkracht, zelfstandige/freelancer, vrije beroepsbeoefenaar, directeur NV of BV, directeur-groootaandeelhouder. Uit deze vraag selecteerden we de werknemers uit de antwoordcategorieën werknemer in vaste of tijdelijke dienst en oproepkracht.

Wanneer respondenten bij eerdere variabelen hebben aangegeven zelfstandige/freelancer of vrije beroepsbeoefenaar te zijn, maar niet in een maatschap of een vennootschap onder firma te zitten, dan kregen ze daarna de vraag: *Bent/was u zelfstandige of werkt/werkte u mee in het gezinsbedrijf of was u freelancer?* De antwoordopties waren: ‘zelfstandige’, ‘werk(te) mee in gezinsbedrijf’ en ‘freelancer’. De antwoordcategorieën ‘zelfstandige’ en ‘freelancer’ hebben we geselecteerd als ‘zelfstandige’. In de uiteindelijke variabele hebben we zo twee categorieën: 0 werknemer en 1 zelfstandige, waarbij zelfstandige staat voor mensen met een eenmanszaak.

Sector

Natuurlijk heeft geen van de zelfstandigen een overheid of een semi-overheid als werkgever, dus deze variabele is toegevoegd als een extra groep van werknemers. De vraag voor deze variabele is: *In welk type organisatie werkt/werkte u (in uw laatste baan)?* De antwoordcategorieën zijn: 1) overheid, semi-overheid en 2) particulier bedrijf. De uiteindelijke variabele is geconstrueerd als werknemer private sector met 1) werknemer private sector en 0) niet werknemer private sector. In combinatie met de dummy ‘zelfstandige’ zijn er dus drie groepen met verschillende ‘werksituatie’. De uiteindelijke groepen zijn daardoor ‘zelfstandige’, ‘werknemer private sector’ en ‘werknemer overheid/semi-overheid’.

Leeftijd

Voor ‘leeftijd’ is er geselecteerd op leeftijden tussen de 18 en 110 jaar. Respondenten jonger dan 18 mogen nog niet stemmen en kunnen daarmee de uitkomsten van de analyses ongewenst beïnvloeden. Waarden hoger dan 110 beschouwen we als niet-realistisch. Na deze selectie is de variabele gecentreerd; dat betekent dat de nul-waarde van de variabele komt te liggen op het oorspronkelijke gemiddelde. Hiermee zijn de coëfficiënten van *leeftijd* makkelijker te interpreteren in de uiteindelijke analyses, want zo hoeven er niet eerst 18 stappen worden genomen voordat de waarde relevant is.

Geslacht

In de uiteindelijke variabele van ‘geslacht’ (in de modellen: man) zijn de categorieën 1) man en 0) vrouw. De categorie ‘anders’ is als missing gecodeerd, omdat deze te klein was om in de analyses mee te nemen.

Migratieachtergrond

‘Migratieachtergrond’ is gecodeerd uit een achtergrondvariabele met de categorieën: ‘autochtone herkomst’, ‘eerste generatie niet-Nederlandse westerse afkomst’, ‘eerste generatie van niet-Nederlandse niet-westerse afkomst’, ‘tweede generatie van niet-Nederlandse westerse afkomst’ en ‘tweede generatie van niet-Nederlandse niet-westerse afkomst’. Om te zorgen dat de categorieën niet te klein worden, zijn ze samengevoegd tot twee dummy’s: ‘eerste generatie’ en ‘tweede generatie’, waardoor ‘geen migratieachtergrond’ de referentiegroep is.

3.3 Analyseopzet

Voor de hypothesetoetsing zijn verschillende binaire logistische modellen geschat, waarbij passieve en actieve politieke participatie de afhankelijke variabelen zijn. Deze modellen schatten de *odds* waarop een respondent meedoet aan beide vormen van politieke participatie. Deze *odds* kunnen we gebruiken om de geschatte proporties in de steekproef te berekenen. Aan de hand van de modellen kunnen we vervolgens de hypothesen toetsen. Om te beoordelen of een waarde significant is hanteren we een significantieniveau van $\alpha=0,01$. Dit significantieniveau houden we redelijk hoog, omdat er niet volledig sprake is van onafhankelijke steekproeftrekking: er kunnen namelijk meerdere personen uit hetzelfde huishouden in de steekproef komen en die personen kunnen elkaars waarden beïnvloeden. Deze schending betekent dat de steekproefomvang wordt overschat en dus er makkelijker

significante resultaten worden gevonden. Het hogere significantieniveau kan hiervoor compenseren. Hier gaan we breder op in, in bijlage 3 onder het kopje assumptie.

Tabel 3 toont de gebruikte modellen; in totaal worden er vier modellen geschat. Bij modellen A en B zijn de onafhankelijke variabelen: leeftijd, geslacht, migratieachtergrond, werksituatie (zelfstandige/werknemer) en opleidingsniveau. In modellen C en D is er geselecteerd op respondenten met een hoog opleidingsniveau en zijn de onafhankelijke variabelen: leeftijd, geslacht, migratieachtergrond en werksituatie. Aan de hand van deze modellen toetsen we de hypothesen als volgt. De hypothesen toetsen we als volgt:

H1) Het effect van opleidingsniveau is sterker voor werknemers dan voor zelfstandigen op passieve politieke participatie.

Voor deze hypothese gebruiken we model A. Deze hypothese beschrijft een versterkend effect van werknemerschap op de invloed van opleidingsniveau op passieve politieke participatie. Werknemerschap is dus een moderator in dit model en dit kan getoetst worden door interactie-variabelen aan het model toe te voegen. Doordat ‘zelfstandige’ een dummy-variabele is en opleidingsniveau ook bestaat uit dummy’s, zijn de interactie-effecten ook dummy’s. Door te kijken of de interactie-effecten significant zijn, kunnen we de hypothese toetsen.

H2) Het effect van opleidingsniveau is sterker voor werknemers dan voor zelfstandigen op actieve politieke participatie.

Hypothese 2 toetsen we op dezelfde manier als hypothese 1 alleen dan uiteraard met model B, waarmee we de proporties van actieve politieke participatie schatten.

H3) Hoogopgeleide werknemers hebben een hogere mate van politieke participatie dan hoogopgeleide zelfstandigen.

Voor deze hypothese kijken we naar model C. Dit model schat de proporties van passieve politieke participatie, waarbij er geselecteerd is op hoogopgeleiden. De hypothese kunnen we dan toetsen door te kijken of de coëfficiënt van zelfstandigen significant negatief is.

H4) Hoogopgeleide werknemers hebben een hogere mate van actieve politieke participatie dan hoogopgeleide zelfstandigen.

Deze hypothese toetsen we weer op dezelfde manier als hypothese 3 alleen dan bij model D waarin de proporties van actieve politieke participatie worden geschat, geselecteerd op hoogopgeleiden.

H5) De invloeden van opleidingsniveau en zelfstandig ondernemerschap zijn groter voor actieve vormen van politieke participatie dan voor passieve vormen.

Met deze hypothese zeggen we dat de coëfficiënten van opleidingsniveau en zelfstandige groter zijn voor actieve politieke participatie significant dan voor passieve politieke participatie. Dit kunnen we beoordelen door terug te keren naar modellen A en B en daar te kijken of de betrouwbaarheidsintervallen van de coëfficiënten van opleidingsniveau en zelfstandige overlappen of niet. Wanneer er geen overlap is, betekent dat de invloeden significant van elkaar verschillen. In dat geval ondersteunen de data de hypothese.

Tabel 1: Overzicht van de binaire regressiemodellen.

Model	Hypothesen	Extra selectie	Afhankelijke variabele
A	1 & 5		Passieve politieke participatie
B	2 & 5		Actieve politieke participatie
C	3	Hoogopgeleid	Passieve politieke participatie
D	4	Hoogopgeleid	Actieve politieke participatie

4. Resultaten

In deze paragraaf bespreken we de uitkomsten van de analyses. Om inzicht te geven in de variabelen en hoe ze onderling samenhangen bespreken we eerst in paragraaf 4.1 de beschrijvende statistieken. In paragraaf 4.2 behandelen we de kwaliteit van de modellen. Tot slot toetsen we in paragraaf 4.3 de hypotheses aan de hand van de uitkomsten van de regressiemodellen. Een uitgebreidere bespreking van de analyses en de modevaluatie is te vinden in bijlagen 2 (analyses) en 3 (modevaluatie).

4.1 Beschrijvende statistieken

Univariate statistieken

Om inzicht te krijgen hoe de afzonderlijke variabelen eruitzien kijken we naar de univariate statistieken. Tabel 2 geeft voor de categorische variabelen de procentuele verdelingen weer en voor de continue variabele ‘leeftijd’ het gemiddelde, standaarddeviatie, minimum, maximum en de kwartielen. Vanwege de selectie op respondenten met een score op elke variabele is de steekproefomvang voor elke variabele gelijk, namelijk $n=2604$.

We zien dat de meerderheid van de respondenten (89,7%) wel aan passieve politieke participatie heeft gedaan tegenover 10,3% die dat niet heeft gedaan. Bij actieve politieke participatie is dat andersom; daar heeft maar 17,8% van de respondenten aan meegedaan tegenover 82,2% niet. Kijken we naar de onafhankelijke variabelen dan hebben de meeste respondenten een hoog opleidingsniveau (48,6%), met 37,7% gemiddeld en 14,1% laag opleidingsniveau. De andere onafhankelijke variabele is ‘werksituatie’: 7,6% van de steekproef bestaat uit zelfstandigen, tegenover 35,1% werknemers die werken bij de overheid of een semi-overheid en 57,3% die als werknemer werken in de private sector. Tot slot zijn er nog de controlevariabelen. De meeste respondenten hebben geen migratieachtergrond (82,7%), tegenover 7,0% eerstegeneratie-immigranten en 10,3% tweedegeneratie-immigranten. Van de steekproef is 53,6% vrouw en 46,4% man. De gemiddelde leeftijd is 49 met standaarddeviatie van 13,9 en de leeftijden lopen uiteen van 19 tot 86.

Tabel 2: Univariate statistieken van de in de analyses opgenomen variabelen: gemiddelde (standaarddeviatie), minimum- en maximumwaarde en totaal aantal respondenten.

Variabele	Gemiddelde (standaard deviatie) ^a	Minimum	Maximum	Kwartielen	n totaal
Passieve pol. Participatie	Niet: 10,3% Wel: 89,7%				2604
Actieve pol. Participatie	Niet: 82,2% Wel: 17,8%				2604
Opleidingsniveau	Laag: 14,1% Gemiddeld: 37,7% Hoog: 48,6%				2604
Werksituatie	Overheidssector: 35,1% Private sector: 57,3% Zelfstandige: 7,6%				2604
Migratieachtergrond	Geen migratieachtergrond: 82,7% Eerste generatie: 7,0% Tweede generatie: 10,3%				2604
Geslacht	Vrouw: 53,6% Man: 46,4%				2604
Leeftijd	-1,94 (13,9) 49*	-31,9 19*	35,1 86*	Q1: 13,7 Q2: 0,1 Q3: 10,1	2604

*Vóór centreren van leeftijd op het gemiddelde

^a Voor categorische variabelen wordt de verdeling in percentages weergegeven.

Bivariate statistieken

Om de verbanden tussen variabelen te ontdekken kijken we naar de bivariate statistieken zoals ze zijn weergegeven in tabel 3. We kijken daarbij vooral naar de verwachte verbanden en eventuele opvallende verbanden. Hiermee kunnen we al vooruitlopen op de uiteindelijke regressiemodellen. Zo kan er sprake zijn van multicollineariteit of kan een verband tussen afhankelijke variabele wegvallen of juist zichtbaar worden in de uiteindelijke regressiemodellen.

We gebruiken drie verschillende maten van associatie, ofwel correlatie (r). Voor de samenhang tussen de dummy-variabelen en de continue variabele gebruiken we de Pearson-correlatie. In dit geval kunnen we de correlatie interpreteren als het verschil in gemiddelden tussen beide categorieën. De correlatie tussen de categorische variabelen (zowel nominale als dummy's) berekenen we met Cramer's V. Bij de correlaties van Cramer's V is het van belang om er rekening mee te houden dat de correlatie geen informatie geeft over de richting van het verband. Tot slot toetsen we de samenhang tussen de dummy-variabelen en de continue

variabele leeftijd met een F -toets, waarvoor we vervolgens de correlatie handmatig uitrekenen

met de formule $r = \sqrt{1 - \frac{n-k}{n-k+(k-1)*F}}$, waarbij n staat voor de steekproef omvang en k =

aantal categorieën van de variabele.

Allereerst kijken we naar de verbanden van de afhankelijke variabelen. Passieve politieke participatie hangt zoals verwacht samen met de onafhankelijke variabelen opleidingsniveau ($r=0,091$, $p<0,001$) en werksituatie ($r=0,088$, $p<0,001$). Daarbij hebben hoogopgeleiden veel meer aan passieve politieke participatie gedaan (94,1%) dan gemiddeld (86,2%) en laagopgeleiden (83,9%). Bij werksituatie doen werknemers uit de overheidssector het meest aan passieve politieke participatie (92,8%), tegenover 87,9% van de werknemers uit de private sector en 92,4% van de zelfstandigen. Hier valt al op dat de verschillen tussen deze groepen niet heel groot is, waardoor het lastiger zou kunnen zijn om in de uiteindelijke modellen significante relaties te zien. Passieve politieke participatie heeft daarnaast significante correlaties met de controlevariabelen migratieachtergrond ($r=0,132$, $p<0,001$) en leeftijd ($r=0,081$, $p<0,001$). Er zijn grote verschillen in passieve politieke participatie op basis van de migratieachtergrond: van de respondenten zonder migratieachtergrond heeft 91,3% passief politiek geparticipeerd, tegenover 76,4% van de eerste generatie en 85,8% van de tweede generatie. Er zijn geen significante verschillen in passieve politieke participatie tussen mannen en vrouwen: $r=0,001$, $p=0,956$.

De andere afhankelijke variabele, actieve politieke participatie, heeft ook significante relaties met beide onafhankelijke variabelen. Wederom zorgt opleidingsniveau ($r=0,124$, $p<0,001$) voor grote verschillen: laag 10,4%, gemiddeld 14,5% en hoog 22,5%. Bij werksituatie zijn de verschillen kleiner ($r=0,071$, $p=0,001$): 20,1% van de werknemers uit de overheidssector, 15,6% van de werknemers uit de private sector en 23,6% van de zelfstandigen. Opvallend is dat er geen significante verschillen zijn tussen de verschillende groepen van migratieachtergrond ($r=0,039$, $p=0,132$). In tegenstelling tot bij passieve politieke participatie heeft actieve politieke participatie wel een significante correlatie: $r=0,073$, $p<0,001$. 20,8% van de mannen doet aan actieve politieke participatie tegenover 15,2% van de vrouwen. Er is tot slot geen significante correlatie met leeftijd: $r=0,017$, $p=0,374$.

Wanneer we kijken naar de onafhankelijke variabelen en de controlevariabelen onderling dan zien we een aantal significante verbanden. Het is daarom van belang om bij de modevaluatie (paragraaf 4.2) goed te controleren voor multicollineariteit. Zo is er een redelijk sterk verband tussen werksituatie en opleidingsniveau ($r=0,138$, $p<0,001$). We zien dat hoogopgeleiden

relatief veel als werknemer bij de overheidssector (43,4%) en als zelfstandige werken (9,1%). Gemiddeld- en laagopgeleiden werken relatief vaker als werknemer in de private sector (respectievelijk 65,1% en 70,2%). Verder heeft werksituatie een significante maar kleine correlatie met migratieachtergrond ($r=0,043$, $p=0,047$) en geslacht ($r=0,152$, $p<0,001$). Ten slotte heeft leeftijd significante correlaties met opleidingsniveau ($F=41,632$, $r=0,176$ en $p<0,001$), werksituatie ($F=32,132$, $r=0,155$, $p<0,001$) en migratieachtergrond ($F=24,114$, $r=0,135$, $p<0,001$).

Tabel 3: Bivariate statistieken van de variabelen.

	1	2	3	4	5	6	7
1 Passieve politieke participatie	-						
2 Actieve politieke participatie	0,091**b	-					
3 Opleidingsniveau	0,142**b	0,124**b	-				
4 Werksituatie	0,088**b	0,071**b	0,138**b	-			
5 Migratieachtergrond	0,132**b	0,039	0,020 ^b	0,043**b	-		
6 Geslacht	0,001 ^b	0,073**b	0,019 ^b	0,152**b	0,043 ^b	-	
7 Leeftijd	0,082**a	0,017 ^a	0,176**c	0,155**c	0,135**c	0,094**a	-

^a is Pearson correlatie; ^b is Cramer's V; ^c op basis van F-toets met Anova
^{*}significant bij $p<0,05$; ^{**}significant bij $p<0,01$; (tweezijdige toets); $n=2604$.

4.2 Modevaluatie

Alvorens we naar de regressiemodellen gaan kijken om uitspraken te doen over de hypotheses, beoordelen we eerst naar de kwaliteit van de modellen. Op die manier kunnen we de resultaten met de juiste nuance interpreteren. Deze modevaluatie bestaat uit vier punten: assumptie van binaire logistische regressie, *outliers* en invloedrijke punten, multicollineariteit en de fit van de modellen. Een meer uitgebreide bespreking van de modevaluatie is te vinden in bijlage 3.

Assumptie

Binaire logistische regressie vereist één assumptie: onafhankelijke steekproeftrekking. Dit betekent dat respondenten onderling geen invloed hebben op de scores van elkaar. Aan deze assumptie wordt niet geheel voldaan, gezien er respondenten uit dezelfde huishoudens in de steekproef kunnen komen. Deze respondenten kunnen elkaar beïnvloeden, bijvoorbeeld doordat de één de ander aanstuurt om te gaan stemmen. Het gevolg van deze assumptieschending is dat de steekproefomvang wordt overschat en dus dat de p -waardes en

de standaardfouten worden onderschat. Daarmee worden de resultaten dus eerder significant. Om hiervoor te controleren gebruiken we een significantieniveau van $\alpha=0,01$ in plaats van het gebruikelijke $\alpha=0,05$.

Outliers en invloedrijke punten

Outliers en invloedrijke punten zijn individuele respondenten die een groot effect op de uitkomsten van de analyses hebben doordat ze extreme of sterk afwijkende scores hebben. Om deze respondenten te vinden kijken we naar twee maten: de DFBETA en de gestandaardiseerde residuen van individuele cases. De DFBETA staat voor de mate waarin een respondent van invloed is op de coëfficiënt van een variabele. De kritieke waarde die we daarvoor hanteren berekenen we met de formule: $\frac{2}{\sqrt{n}}$. Daarnaast kijken we dus naar de gestandaardiseerde residuen, waarbij we kritieke waardes hanteren van 3 en -3. Een te hoog gestandaardiseerd residu staat voor dat de kans van de afhankelijke variabele is onderschat. Een te laag gestandaardiseerd residu betekent dat de kans is overschat.

Wanneer we de modellen schatten en de geschatte waarden vergelijken met de kritieke waarden, dan zijn er redelijk veel invloedrijke punten. Model A, met als afhankelijke variabele passieve politieke participatie, heeft 79 respondenten met een te hoog gestandaardiseerd residu, dus 79 cases waarvan hun kans op passieve politieke participatie is overschat. Als we inzoomen op deze cases dan zien we dat dit respondenten zijn die niet hebben geparticipeerd, maar wel de kenmerken hebben die sterk correleren met passieve politieke participatie: veelal hoogopgeleid, werkend in de overheidssector en geen migratieachtergrond. Daarnaast zien we dat in model A de coëfficiënt van zelfstandige omlaag wordt gebracht door een aantal invloedrijke punten.

Model B, met als afhankelijke variabele actieve politieke participatie, heeft ook te maken met cases die een te veel afwijkende gestandaardiseerd residu hebben, maar nu juist te laag. Dat betekent dat de kans op actieve politieke participatie is overschat. Kijken we naar deze cases, dan zien we dat ze vaak hoogopgeleid zijn, veelal werken bij de overheid en geen migratieachtergrond hebben. Het lijkt er daarom op dat deze groep dezelfde groep is als die van model A. Model B heeft daarnaast ook te maken met te hoge DFBETA's voor de coëfficiënt van zelfstandige, wat betekent dat deze coëfficiënt omhoog is getrokken door een aantal invloedrijke punten.

Model C is het model waarbij de kans op passieve politieke participatie wordt geschat en geselecteerd is op hoogopgeleiden. Deze kent net als model A ook te lage gestandaardiseerde

residuen (-7,25). Daarnaast heeft dit model ook problemen met verschillende DFBETA's, namelijk die van zelfstandige, eerste generatie en tweede generatie. Voor elk van deze variabelen geldt dat de waarde te laag is, dus dat de coëfficiënten van elk van deze variabelen omlaag worden getrokken door invloedrijke punten.

Model D, waarbij de kans op actieve politieke participatie wordt geschat en geselecteerd is op hoogopgeleiden, heeft geen problemen met invloedrijke punten.

Echter, er is geen reden om aan te nemen dat deze invloedrijke punten het gevolg zijn van niet-realistische of foutieve waarden. Daarom blijven deze invloedrijke punten in de uiteindelijke steekproef.

Multicollineariteit

Multicollineariteit betekent dat de verklarende variabelen te veel overlap hebben in hun verklarende kracht, waardoor ze in het model een minder sterk effect laten zien. Het gevolg is dat er grotere standaardfouten ontstaan en coëfficiënten minder snel significant zijn. Om te controleren voor multicollineariteit kijken we naar de VIF-scores van de variabelen bij het schatten van lineaire regressiemodellen. Deze scores mogen niet hoger zijn dan de kritieke waarde van 4. Maar dit is niet het geval, dus geen van de modellen is er dus sprake van multicollineariteit.

Fit van de modellen

De *fit* van de logistische modellen betekent in hoeverre de modellen in staat zijn goede voorspellingen te doen van de waarde van de afhankelijke variabele. Daarvoor kijken we naar drie punten: de *deviance*, de *Hosmer and Lemeshow test* en de classificatiemodellen. De *deviance* staat voor de slechtheid van de fit van een model, dus een hogere *deviance* betekent een slechtere fit. Je zou het kunnen vertalen als 'afwijking van het perfecte model'. Per stap in de modellen kijken we of de *deviance* significant omlaag is gegaan. De *Hosmer and Lemeshow test* deelt de steekproef op basis van voorspelde kansen in tien groepen en kijkt dan in hoeverre de juiste proportie van de afhankelijke variabele is geschat. Deze toetst dan de nulhypothese dat het model een slechte voorspeller is, dus een hogere *p*-waarde betekent een beter model. Tot slot kijken we naar de classificatietabellen van de nul-modellen en de uiteindelijke modellen. De nul-modellen zijn de modellen waarin alleen een constante is opgenomen, ofwel het gemiddelde. De classificatiemodellen schatten de proporties van de afhankelijke variabele op basis van die modellen. Wordt er een proportie geschat die dichter

bij de daadwerkelijke proportie ligt, dan betekent dat dat model dus een betere voorspeller van de afhankelijke variabele is.

Model A: Passieve politieke participatie

Voor dit model kijken we eerst naar de *deviance* zoals ook weergegeven in tabel 3. De *deviance* van het eerste model waarin de controlevariabelen leeftijd, geslacht, migratieachtergrond en de categorie werknemers uit de private sector zijn opgenomen, is 1659,01. Aan de X^2 -toets is te zien dat de toevoeging van deze variabelen voor een significante verbetering van de *fit* zorgt: $X^2=67,173$ met $p<0,001$. In het tweede blok zijn daar de onafhankelijke variabelen zelfstandige en opleidingsniveau aan toegevoegd. We zien dan dat de *deviance* afneemt naar 1601,34 ($X^2=57,658$ $p<0,001$). De daling in *deviance* is daarmee niet heel groot, maar wel significant. In blok 3 komen daar nog de interactievariabelen bij, maar dat heeft geen merkbare invloed op de *deviance*; deze blijft 1601,34 ($X^2=0,040$, $p=0,980$). Vervolgens kijken we voor dit model naar de *Hosmer and Lemeshow test*. Deze toont een X^2 -waarde van 20,332 met $p=0,009$. Dat betekent dat het resultaat significant is, waarmee deze test zegt dat het model geen goede *fit* heeft. De classificatietabellen voorspellen bijna hetzelfde voor het nul-model als voor het uiteindelijke model. Door de hoge kans op passieve politieke participatie komen slechts twee respondenten van de steekproef over de *cut-off value* van 0,50.

Over dit model kunnen we concluderen dat de *fit* van het model niet heel goed is, waardoor het model waarschijnlijk niet een hele sterke voorspeller van passieve politieke participatie is. Weliswaar is de *deviance* door toevoeging van de controle- en de onafhankelijke variabelen wel significant afgenomen, dus het model is zeker een betere voorspeller dan het nul-model. Toch moeten we hierdoor voorzichtig zijn bij het interpreteren van de resultaten.

Model B: Actieve politieke participatie

De *deviance* van dit model is 2403,83 ($X^2=33,759$, $p<0,001$) wanneer alleen de controlevariabelen zijn toegevoegd. Wanneer de onafhankelijke variabelen daaraan toegevoegd worden daalt die *deviance* naar 2367,58 ($X^2=36,429$, $p<0,001$). Wederom laat toevoeging van de onafhankelijke variabelen niet een hele grote daling in de *deviance* zien. Wanneer ook de interactievariabelen zijn toegevoegd daalt de *deviance* verder naar 2365,91 met $X^2=1,665$ en $p=0,435$. Weliswaar is deze daling nog niet significant, wel zien we dat de interactievariabelen voor meer verbetering van de *fit* van het model zorgen dan het geval was bij model A. Kijken we naar de *Hosmer Lemeshow test*, dan zien we dat daar een redelijke

score uit komt: $X^2=7,119$, $p=0,524$. De p -waarde is redelijk hoog, wat betekent dat het model een betere voorspeller van de proportie van de afhankelijke variabele is. Tot slot kijken we weer naar de classificatietabellen. Hier treedt hetzelfde probleem op als bij model A het geval was, alleen dan in de andere richting: geen van de respondenten heeft een kans hoger dan de *cut-off value* 0,5 om mee te doen aan actieve politieke participatie, waardoor de classificatietabel van het uiteindelijke model gelijk is aan die van het nul-model.

Concluderend kunnen we zeggen dat dit model een redelijke voorspeller van de proporties van actieve politieke participatie lijkt.

Model C: Passieve politieke participatie geselecteerd op hoogopgeleiden

De modellen C en D hebben een stap minder, doordat opleidingsniveau en dus ook de interactievariabelen van opleidingsniveau en zelfstandige niet in de modellen zijn opgenomen.

In blok 1 zijn dezelfde controlevariabelen toegevoegd, daarmee is de *deviance* 527,896 ($X^2=41,368$ en $p<0,001$). Dit blok is dus een significante verbetering ten opzichte van het nul-model waar alleen de constante in is opgenomen. In blok 2 is alleen zelfstandige nog toegevoegd; daarmee daalt de *deviance* miniem naar 527,857 ($X^2=0,039$ en $p=0,844$). Dit duidt erop dat de toevoeging van zelfstandige het model niet een betere voorspeller maakt van de proporties van passieve politieke participatie. De *Hosmer and Lemeshow test* heeft weliswaar geen significant resultaat, de p -waarde is wel redelijk laag wat betekent dat de *fit* van het model niet zo goed is. Wederom geven de classificatietabellen geen verschil in resultaat tussen het nul-model en het uiteindelijke model.

Net zoals model A heeft dit model dus geen hele goede *fit*, waardoor de voorspelde proporties waarschijnlijk minder nauwkeurig zijn.

Model D: Actieve politieke participatie geselecteerd op hoogopgeleiden

In het eerste blok is de *deviance* 1326,38 met $X^2=20,974$ en $p<0,001$. Het model met de controlevariabelen kent dus een significant betere *fit* dan het model met alleen een constante. In blok 2 heeft wederom de toevoeging van zelfstandige als voorspeller maar weinig invloed op de *deviance*: 1324,85 ($X^2=1,532$ en $p=0,216$). De *Hosmer and Lemeshow test* geeft een resultaat van ($X^2=9,368$, $p=0,312$), waarmee de nulhypothese dat het model een goede *fit* heeft niet wordt verworpen. De p -waarde is desondanks niet heel hoog. Tot slot zijn ook bij dit

model de classificatietabellen van het nul-model en het uiteindelijke model gelijk doordat de verwachte proporties niet hoger zijn dan 0,5.

Uiteindelijk is het model dus een redelijke voorspeller van de proporties van actieve politieke participatie, maar de toevoeging van de onafhankelijke variabele zorgt niet voor veel nauwkeurigere voorspellingen.

4.3 Hypothesetoetsing

Voor de hypothesetoetsing toetsen we de parameters van de binaire logistische modellen in tabellen 4, 5, 6 en 7. De tabellen geven de parameters van de geschatte proporties van de afhankelijke variabelen: voor tabel 4 en 6 is dat passieve politieke participatie en voor 5 en 7 is dat actieve politieke participatie. Deze modellen zijn stapsgewijs opgebouwd, zodat zichtbaar is wat de invloeden van de toevoegingen van extra variabelen aan het model verandert. In stap 1 voegen we de controlevariabelen toe, in stap 2 de onafhankelijke variabelen en eventueel in stap 3 de interactievariabelen. Elk van de modellen heeft daarnaast een constante; een negatieve constante staat voor een basiskans van lager dan 0,5 en een positieve constante staat voor een basiskans hoger dan 0,5. Met basiskans wordt bedoeld de kans die iemand heeft die op alle variabelen nul scoort en dus behoort tot alle referentiecategorieën. De coëfficiënten (b) staan voor de *log-odds* van de geschatte kansen van de afhankelijke variabelen. Hierdoor zijn de coëfficiënten op zichzelf lastig te interpreteren. Daarom rekenen we voor een aantal hypothetische respondenten de voorspelde kans uit. Dit doen we door de regressievergelijking in te vullen en die uitkomst in de volgende formule toe te passen:

$$\hat{p} = \frac{e^y}{1 + e^y}$$

Daarbij staat \hat{p} voor de geschatte kans en is y de uitkomst van de regressievergelijking in de macht van e . Naast de coëfficiënten zijn ook de odds-ratio's gegeven; dit staat voor de waarde waarmee de odds worden vermenigvuldigd wanneer de onafhankelijke variabele met één eenheid stijgt. Aan de hand van deze parameters kunnen we uitspraken doen over de hypothesen. Voor de rest van deze paragraaf nemen we de hypothesen als leidraad voor de structuur.

H1) Het effect van opleidingsniveau is sterker voor werknemers dan voor zelfstandigen op passieve politieke participatie.

Deze hypothese toetsen we met de coëfficiënten van de interactievariabele van model A in tabel 4. Daar is te zien dat de interactievariabelen in het uiteindelijke model zeer kleine coëfficiënten hebben en zeer hoge p -waardes: laagopgeleid*zelfstandige ($b=-0,077$, odds-ratio=0,926 en $p=0,920$) en hoogopgeleid*zelfstandige ($b=0,080$, odds-ratio=1,083, $p=0,900$). Dat betekent dat er eigenlijk geen effect van de interactievariabelen op de afhankelijke variabele is gemeten. De interacties zijn niet significant en hebben maar zeer kleine coëfficiënten. Dat sluit aan bij de waarneming tijdens de modevaluatie waar we zagen dat toevoeging van de interacties geen zichtbare verandering had op de fit van het model.

Om het model beter te kunnen interpreteren, zijn voor een aantal respondenten de geschatte kans uitgerekend. We vergelijken een zelfstandige met laag ($\hat{p}=0,858$) en met hoog opleidingsniveau ($\hat{p}=0,961$), een werknemer uit de private sector met laag ($\hat{p}=0,842$) en hoog opleidingsniveau ($\hat{p}=0,949$) en tot slot werknemers uit de overheidssector met laag ($\hat{p}=0,884$) en hoog opleidingsniveau ($\hat{p}=0,964$) (allen vrouw, zonder migratieachtergrond en met gemiddelde leeftijd). Ook aan deze coëfficiënten is te zien dat opleidingsniveau niet voor grotere verschillen in passieve politieke participatie zorgt bij werknemers dan bij zelfstandigen. Dit betekent voor de hypothese dat deze wordt verworpen: het effect van opleidingsniveau is niet sterker voor werknemers dan voor zelfstandigen op passieve politieke participatie.

Wanneer we de andere coëfficiënten van het model analyseren, dan is dit resultaat niet verassend: de coëfficiënt die het verschil in de kans op passieve politieke participatie tussen zelfstandigen en werknemers weergeeft is namelijk ook niet significant ($b=-0,157$, odds-ratio=0,855, $p=0,720$). Daarnaast is ook de coëfficiënt van het verband van laagopgeleid met passieve politieke participatie niet significant ($b=-0,295$, odds-ratio=0,744 en $p=0,103$), maar die van het verband van hoogopgeleid is dat wel ($b=0,958$, odds-ratio=2,608, $p<0,001$). Bovendien is er een positieve constante ($b=2,325$, $p<0,001$) en zijn er significante controlevariabelen leeftijd ($b=0,023$, odds-ratio=1,023, $p<0,001$) en eerste generatie immigranten ($b=-1,256$, odds-ratio=0,285, $p<0,001$). De effecten van geslacht ($b=0,023$, odds-ratio=1,023, $p=0,864$), tweede generatie ($b=-0,435$, odds-ratio=0,647 en $p=0,028$) en werknemers uit de private sector ($b=-0,355$, odds-ratio=0,702, $p=0,024$) zijn niet significant, al hebben die laatste twee wel redelijk lage p -waardes.

H2) Het effect van opleidingsniveau is sterker voor werknemers dan voor zelfstandigen op actieve politieke participatie.

Hypothese 2 toetsen we met model B, waarmee de proporties van actieve politieke participatie worden geschat. De parameters van dit model zijn te vinden in tabel 5. In tabel 5 is te zien dat de interactievariabelen hogere coëfficiënten en lagere p -waardes hebben dan bij passieve politieke participatie: zelfstandige*laag ($b=0,264$, odds-ratio=1,302, $p=0,723$) en zelfstandige*hoog ($b=0,540$, odds-ratio=1,715, $p=0,214$). Bovendien zijn de coëfficiënten van de verbanden tussen de afhankelijke variabele met zelfstandige en met opleidingsniveau na toevoeging van de interactie-effecten een klein beetje veranderd. De interactie-effecten hebben in dit model dus wel enige invloed op de afhankelijke variabele, maar die invloed is niet significant. Dat zagen we ook terug in de modevaluatie, waar de fit van het model enigszins verbeterd werd. Opvallend is wel dat het interactie-effect van ‘zelfstandige’ met ‘laagopgeleid’ positief is terwijl de beide coëfficiënten van de effecten van laagopgeleid ($b=-0,421$, odds-ratio=0,657, $p=0,039$) en zelfstandige negatief zijn ($b=-0,205$, odds-ratio=0,815, $p=0,592$).

Rekenen we voor de dit model de geschatte kansen uit dan zijn die voor vrouwen zonder migratieachtergrond en met gemiddelde leeftijd: laag ($\hat{p}=0,099$) en hoogopgeleide ($\hat{p}=0,258$) zelfstandige, laag ($\hat{p}=0,073$) en hoogopgeleide ($\hat{p}=0,160$) werknemer uit de private sector en tot slot laag ($\hat{p}=0,099$) en hoogopgeleide ($\hat{p}=0,199$) werknemer uit de overheidssector. Gekeken naar deze voorspelde kansen dan valt op dat de kansen van zelfstandigen op actieve politieke participatie verder uit elkaar liggen dan die van de werknemers. Gezien interactie-effecten niet significant zijn, verwerpen we hypothese 2.

Aan de andere coëfficiënten in het model valt op dat de constante negatief is, wat betekent dat de basiskans lager dan 0,5 is. Verder is te zien dat in dit model andere significante controlevariabelen optreden dan in het model waarmee de proporties van passieve politieke participatie worden geschat. De invloed van geslacht heeft nu namelijk wel een significante coëfficiënt ($b=0,426$, odds-ratio=1,532 en $p<0,001$) en de coëfficiënt die het verschil tussen de tweede generatie immigranten ten opzichte van mensen zonder migratieachtergrond toont, heeft een vrij lage p -waarde ($b=0,338$, odds-ratio=1,402 en $p=0,038$). De verbanden van leeftijd ($b=0,004$, odds-ratio=1,004, $p=0,289$) en eerste generatie ($b=0,134$, odds-ratio 1,143 en $p=0,508$) hebben lage coëfficiënten en geen lage p -waardes. We nemen dus verschillende verbanden waar voor actieve en passieve politieke participatie.

H3) Hoogopgeleide werknemers hebben een hogere mate van passieve politieke participatie dan hoogopgeleide zelfstandigen.

Voor deze hypothese richten we ons op model C in tabel 6. Dit model schat passieve politieke participatie op basis van leeftijd, geslacht, migratieachtergrond en werksituatie. Daarnaast is er geselecteerd op hoogopgeleiden. De hypothese wordt getoetst met het feit of de coëfficiënt van de variabele zelfstandige significant negatief is. Weliswaar is de coëfficiënt in de goede richting, met $b=-0,101$, odds-ratio=0,904 en $p=0,843$ is de variabele verre van significant, waarmee we de hypothese kunnen verwerpen. Daar komt nog bij dat de coëfficiënt als referentiegroep werknemers uit de overheidssector gebruikt, terwijl de coëfficiënt die het verschil toont met werknemers uit de private sector groter is dan die van zelfstandige ($b=-0,279$, odds-ratio=0,756, $p=0,303$). Dat betekent dat hoogopgeleide werknemers uit de private sector een lagere mate van passieve politieke participatie dan hoogopgeleide zelfstandigen. Wanneer we de kansen uitrekenen van het model (van een vrouw met gemiddelde leeftijd en zonder migratieachtergrond), dan zien we dat gegeven uitgedrukt in geschatte proporties: werknemers uit de overheidssector ($\hat{p}=0,970$), werknemers uit de private sector ($\hat{p}=0,935$) en zelfstandigen ($\hat{p}=0,970$).

Wanneer we model C vergelijken met model A, dan valt op dat de coëfficiënten die de invloed van migratieachtergrond tonen sterk verschillen. Dat zou kunnen betekenen dat het effect van migratieachtergrond op de passieve politieke participatie van hoogopgeleiden anders is dan op die van gemiddeld en laagopgeleiden. In model A zijn de coëfficiënten namelijk groter: eerste generatie ($b=-1,256$, odds-ratio=0,285 en $p<0,001$) en tweede generatie ($b=-0,434$, odds-ratio=0,648, $p=0,028$) tegenover model C met eerste generatie ($b=-1,736$, odds-ratio=0,176 en $p<0,001$) en tweede generatie ($b=-0,882$, odds-ratio=0,414, $p=0,007$). Het lijkt er daarmee op dat tussen hoogopgeleiden met verschillende migratieachtergronden de verschillen tussen passieve politieke participatie groter zijn dan voor gemiddeld en lager opgeleiden.

H4) Hoogopgeleide werknemers hebben een hogere mate van actieve politieke participatie dan hoogopgeleide zelfstandigen.

Voor deze hypothese gebruiken we het laatste model (tabel 7). Dit model schat de proporties van actieve politieke participatie. In dit model kijken we weer naar de coëfficiënt die het verschil tussen zelfstandigen en werknemers van de overheid weergeeft, om de hypothese te kunnen toetsen. Die coëfficiënt is in het uiteindelijke model: $b=0,286$, odds-ratio=1,332, $p=0,210$. Dit resultaat is verassend; niet alleen is deze coëfficiënt niet significant, maar ook is deze in tegengestelde richting van wat verwacht. Dit zou betekenen dat hoog opgeleide zelfstandigen eerder een hogere kans hebben op actieve politieke participatie dan hoog

opgeleide werknemers uit de overheidssector. Gezien de variabele van werknemers uit de private sector een negatieve coëfficiënt heeft ($b=-0,304$, odds-ratio= $0,738$, $p=0,045$), hebben werknemers uit de private sector nog een lagere geschatte kans dan werknemers uit de overheidssector. De hypothese kunnen we daarmee verwerpen. Berekenen we de kansen van deze drie groepen (van een vrouw met gemiddelde leeftijd en zonder migratieachtergrond) dan zien die er zo uit: zelfstandigen ($\hat{p}=0,258$), werknemers uit de private sector ($\hat{p}=0,162$) en werknemers uit de overheidssector ($\hat{p}=0,207$). Ook al zijn de coëfficiënten niet negatief, toch zijn de verschillen redelijk groot.

Verder valt op aan het model dat geslacht voor redelijk grote verschillen zorgt in de geschatte proporties van actieve politieke participatie ($b=0,434$, odds-ratio= $1,543$, $p=0,002$).

Omgerekend naar kansen betekent dat mannen een significant hogere kans op actieve politieke participatie hebben dan vrouwen, respectievelijk $\hat{p}=0,288$ tegenover $\hat{p}=0,207$.

H5) De invloed van opleidingsniveau en zelfstandig ondernemerschap is groter op actieve vormen van politieke participatie dan op passieve vormen.

Tot slot vergelijken we voor deze hypothese de betrouwbaarheidsintervallen van de odds-ratio's van de variabelen zelfstandige en opleidingsniveau tussen de modellen A en B (tabellen 4 en 5). Mochten de odds-ratio's van model B buiten de betrouwbaarheidsintervallen van model A vallen, dan is er sprake van significante verschillen in grootte van de effecten. Gezien een odds-ratio van 1 ervoor staat dat de geschatte kans gelijk blijft, verwerpen we de hypothese wanneer de odds-ratio van B verder van 1 is verwijderd dan de grens van het betrouwbaarheidsinterval van model A. De odds-ratio's van de interactievariabelen nemen we niet mee gezien die zeer grote standaardfouten hebben en daardoor ook een zeer ruim betrouwbaarheidsinterval. Het gehanteerde betrouwbaarheidsinterval is 98%, want het betrouwbaarheidsinterval is tweezijdig met $\alpha=0,01$.

In model A zijn de betrouwbaarheidsintervallen van opleidingsniveau als volgt: laag (0,488 tot 1,134) en hoog (1,795 tot 3,789). Die van zelfstandige loopt van 0,309 tot 2,363. Kijken we naar de odds-ratio's van model B, dan zien we dat alleen die van hoog opleidingsniveau buiten het betrouwbaarheidsinterval van model A valt (odds-ratio= $1,575$). De odds-ratio van laag opleidingsniveau is 0,657 en die van zelfstandige is 0,815. De odds-ratio van hoog opleidingsniveau in model B is dus hoger dan het betrouwbaarheidsinterval van model A, wat betekent dat het effect van hoog opleidingsniveau significant groter is op actieve politieke participatie dan op passieve politieke participatie. Voor zowel zelfstandige als laag

opleidingsniveau is dat niet het geval. De hypothese wordt dus door de data voor een gedeelte ondersteund.

J.G. DE VRIES – POLITICAL PARTICIPATION IN THE LIGHT OF THE GOALFRAMING THEORY

Tabel 4: Parameterschattingen van de binaire logistische regressiemodellen met afhankelijke variabele passieve politieke participatie

Model A	Model 1			Model 2			Model 3			Bhi (98%) odds-ratio links	Bhi (98%) Odds-ratio rechts
	<i>b</i> (<i>se</i>)	Odds-ratio	<i>p</i>	<i>b</i> (<i>se</i>)	Odds-ratio	<i>p</i>	<i>b</i> (<i>se</i>)	Odds-ratio	<i>p</i>		
Constante	2,712 (0,133)		<0,001	2,324 (0,163)		<0,001	2,325 (0,164)		<0,001		
Leeftijd	0,017 (0,005)	1,017	<0,001	0,023 (0,005)	1,023	<0,001	0,023 (0,005)	1,023	<0,001	1,011	1,035
Geslacht (1=man, 0=vrouw)	0,034 (0,134)	1,034	0,802	0,023 (0,135)	1,023	0,867	0,023 (0,135)	1,023	0,864	0,748	1,400
Eerste generatie <i>ref.=geen migratieachtergrond</i>	-1,153 (0,193)	0,316	<0,001	-1,256 (0,198)	0,285	<0,001	-1,256 (0,198)	0,285	<0,001	0,180	0,452
Tweede generatie	-0,444 (0,194)	0,641	0,022	-0,434 (0,198)	0,648	0,028	-0,435 (0,198)	0,647	0,028	0,409	1,025
Werknemers private sector <i>Ref: werknemer overheidssector</i>	-0,527 (0,144)	0,590	<0,001	-0,353 (0,157)	0,702	0,025	-0,355 (0,157)	0,702	0,024	0,486	1,012
Zelfstandigen				-0,142 (0,304)	0,868	0,641	-0,157 (0,437)	0,855	0,720	0,309	2,363
Laag opleidingsniveau <i>Ref: gemiddeld opleidingsniveau</i>				-0,300 (0,176)	0,741	0,089	-0,295 (0,181)	0,744	0,103	0,488	1,134
Hoog opleidingsniveau				0,964 (0,155)	2,622	<0,001	0,958 (0,161)	2,608	<0,001	1,795	3,789
Interacties											
Zelfstandige*laag opleidingsniveau							-0,077 (0,765)	0,926	0,920	0,156	5,493
Zelfstandige*hoog opleidingsniveau							0,080 (0,636)	1,083	0,900	0,247	4,757
<i>Deviance (-2log)</i>	1659,013			1601,354			1601,314				
<i>X²-toets</i>	67,173		<0,001	57,658		<0,001	0,040		0,980		
<i>N</i>	2604			2604			2604				

J.G. DE VRIES – POLITICAL PARTICIPATION IN THE LIGHT OF THE GOALFRAMING THEORY

Tabel 5: Parameterschattingen van de binaire logistische regressiemodellen met afhankelijke variabele actieve politieke participatie.

Model B	Model 1			Model 2			Model 3			Bhi (98%) odds-ratio links	Bhi (98%) odds-ratio rechts
	<i>b</i> (<i>se</i>)	Odds-ratio	<i>p</i>	<i>b</i> (<i>se</i>)	Odds-ratio	<i>p</i>	<i>b</i> (<i>se</i>)	Odds-ratio	<i>p</i>		
Constante	-1,560 (0,090)		<0,001	-1,873 (0,128)		<0,001	-1,844 (0,129)		<0,001		
Leeftijd	0,001 (0,004)	1,001	0,850	0,004 (0,004)	1,004	0,273	0,004 (0,004)	1,004	0,289	0,995	1,013
Geslacht (1=man, 0=vrouw)	0,435 (0,106)	1,546	<0,001	0,420 (0,107)	1,521	<0,001	0,426 (0,107)	1,532	<0,001	1,194	1,964
Eerste generatie <i>ref.: geen migratieachtergrond</i>	0,162 (0,200)	1,176	0,417	0,127 (0,202)	1,136	0,528	0,134 (0,202)	1,143	0,508	0,715	1,828
Tweede generatie	0,310 (0,161)	1,363	0,055	0,335 (0,163)	1,397	0,039	0,338 (0,163)	1,402	0,038	0,960	2,047
Werknemers private sector <i>Ref: werknemer overheidssector</i>	-0,425 (0,106)	0,654	<0,001	-0,261 (0,115)	0,770	0,023	-0,270 (0,115)	0,763	0,019	0,583	0,998
Zelfstandigen				0,179 (0,189)	1,197	0,343	-0,205 (0,383)	0,815	0,592	0,334	1,986
Laag opleidingsniveau <i>Ref: gemiddeld opleidingsniveau</i>				-0,404 (0,196)	0,668	0,040	-0,421 (0,204)	0,657	0,039	0,409	1,055
Hoog opleidingsniveau				0,500 (0,116)	1,649	<0,001	0,454 (0,121)	1,575	<0,001	1,189	2,086
Interacties											
Zelfstandige*laag opleidingsniveau							0,264 (0,764)	1,302	0,723	0,230	7,384
Zelfstandige*hoog opleidingsniveau							0,540 (0,434)	1,715	0,214	0,625	4,710
<i>Deviance (-2log)</i>	2403,826			2367,577			2365,91				
<i>X²-toets</i>	33,759		<0,001	36,249		<0,001	1,665		0,435		
N	2604			2604			2604				

Tabel 6: Parameterschattingen van de binaire logistische regressiemodellen met afhankelijke variabele passieve politieke participatie waarbij er geselecteerd is op hoogopgeleiden.

Model C	Model 1			Model 2		
	<i>b</i> (<i>se</i>)	Odds-ratio	<i>p</i>	<i>b</i> (<i>se</i>)	Odds-ratio	<i>p</i>
Constante	3,446 (0,238)		<0,001	3,463 (0,254)		<0,001
Leeftijd	0,029 (0,010)	1,029	0,004	0,029 (0,010)	1,030	0,004
Geslacht (1=man, 0=vrouw)	0,002 (0,255)	1,002	0,994	0,002 (0,255)	1,002	0,993
Eerste generatie <i>ref.=geen migratieachtergrond</i>	-1,734 (0,313)	0,177	<0,001	-1,736 (0,313)	0,176	<0,001
Tweede generatie	-0,884 (0,328)	0,413	0,007	-0,882 (0,328)	0,414	0,007
Werknemers private sector <i>Ref: werknemer overheidssector</i>	-0,263 (0,258)	0,769	0,308	-0,279 (0,271)	0,756	0,303
Zelfstandigen				-0,101 (0,510)	0,904	0,843
<i>Deviance (-2log)</i>	527,896			527,857		
<i>X²-toets</i>	41,368		<0,001	0,039		0,844
<i>N</i>	1265			1265		

Tabel 7: Parameterschattingen van de binaire logistische regressiemodellen met afhankelijke variabele passieve politieke participatie waarbij er geselecteerd is op hoogopgeleiden.

Model D	Model 1			Model 2		
	<i>b</i> (<i>se</i>)	Odds-ratio	<i>p</i>	<i>b</i> (<i>se</i>)	Odds-ratio	<i>p</i>
Constante	-1,288 (0,111)		<0,001	-1,341 (0,120)		<0,001
Leeftijd	0,009 (0,005)	1,009	0,078	0,009 (0,005)	1,009	0,100
Geslacht (1=man, 0=vrouw)	0,434 (0,142)	1,544	0,002	0,434 (0,142)	1,543	0,002
Eerste generatie <i>ref.=geen migratieachtergrond</i>	0,146 (0,254)	1,157	0,565	0,153 (0,254)	1,166	0,546
Tweede generatie	0,121 (0,226)	1,129	0,592	0,117 (0,226)	1,124	0,604
Werknemers private sector <i>Ref: werknemer overheidssector</i>	-0,354 (0,145)	0,702	0,015	-0,304 (0,151)	0,738	0,045
Zelfstandigen				0,286 (0,229)	1,332	0,210
<i>Deviance (-2log)</i>	1326,383			1324,851		
<i>X²-toets</i>	20,974		<0,001	1,532		0,216
<i>N</i>	1265			1265		

5 Conclusie & discussie

In dit hoofdstuk proberen we de onderzoeksvraag te beantwoorden, reflecteren we op de beperkingen van het onderzoek en komen we met aanbevelingen voor vervolgonderzoek.

5.1 Conclusie

Het doel van dit onderzoek was om meer inzicht te krijgen in welke factoren van invloed zijn op het politieke gedrag van mensen. Daarbij probeerden we verschillen in goalframes aan te tonen die ontstaan door verschillen in ondernemerschap/werknemerschap en opleidingsniveau. De onderzoeksvraag die we hebben onderzocht was: *In hoeverre kunnen we politieke participatie verklaren door verschillen in ondernemerschap of werknemerschap en opleidingsniveau?* Daarbij onderzochten we ook de vraag: *Zijn er daarbij verschillen tussen actieve en passieve vormen van politieke participatie?* Om deze vragen te beantwoorden hebben we binaire logistische regressiemodellen geschat met data van het LISS-Panel uit 2022. Daarbij vergeleken we zelfstandigen en werknemers op basis van hun opleidingsniveau in hun kans op actieve en passieve politieke participatie.

In de theorie zijn de *normative* en de *gain goalframes* naast elkaar gelegd. Met de *gain goalframe* verwachtten we dat zelfstandigen een hogere politieke participatie hebben. Allereerst omdat zelfstandigen grotere persoonlijke belangen ervaren dan werknemers. Voor werknemers zijn namelijk veel regelingen al vastgelegd in sociale akkoorden, terwijl voor zelfstandigen die vaak door de politiek worden bepaald of in ieder geval bediscussieerd. Bovendien zijn zelfstandigen afhankelijker van een economisch welvarende omgeving (Rotolo & Wilson, 2006), dus zijn ze ook meer gemotiveerd om voor de belangen van die omgeving op te komen (Mencken, Smith & Tolbert, 2020). Daarnaast hebben zelfstandigen een hogere interne politieke *self-efficacy*, doordat zij meer vaardigheden ontwikkelen die relevant zijn voor politieke participatie (Wilson & Musick, 1997). We verwachtten dus dat voor zelfstandigen de *gain goalframe* meer dominant zou zijn.

Met de *normative goalframe* richtten we ons op verschillen in de invloed van opleidingsniveau. Opleidingsniveau zorgt namelijk voor meer politieke participatie (Milligan, Moretti & Oreopoulos, 2004). We redeneren dat met de *normative goalframe* hoogopgeleide werknemers meer prikkels krijgen vanuit hun sociale omgeving, omdat zij vaker werken met andere hoogopgeleiden die het individu kunnen prikkelen om te gaan participeren.

Zelfstandigen werken minder vaak met mensen met hetzelfde opleidingsniveau, omdat ze

geen collega's en leidinggevende hebben, dus worden ze ook niet door dit deel van het netwerk beïnvloed om te gaan participeren. Daarom verwachtten we dat bij werknemers de *normative goalframe* meer dominant is.

Echter, uit de resultaten van het onderzoek komt geen ondersteuning voor verschillen in goalframes tussen werknemers en zelfstandigen. Allereerst is er geen verschil in de kans op politieke participatie tussen werknemers en zelfstandigen, waardoor er geen indicatie is voor verschillen in de mate waarin de *gain goalframe* een rol speelt tussen werknemers en zelfstandigen. Daarnaast is er geen ondersteuning voor verschillen in invloed van opleidingsniveau tussen werknemers en zelfstandigen. Dus tonen we ook geen verschillen in politieke participatie aan die we zouden verwachten wanneer de normatieve goalframe meer dominant zou zijn bij werknemers dan bij zelfstandigen. In het geheel is er dus geen indicatie dat er verschillen zijn in de mate waarin de *normative* en de *gain goalframe* tussen groepen optreden.

Daarnaast redeneerden we dat de invloed van ondernemerschap en opleidingsniveau verschillend zou zijn voor passieve en actieve politieke participatie. Passieve politieke participatie bestaat namelijk uit vormen van politieke participatie die weinig inzet vragen en minder zichtbaar zijn. Daarentegen vragen vormen van actieve politieke participatie meer inzet en is individuele deelname daarvan meer zichtbaar.

Hoewel we in de resultaten voor beide vormen van politieke participatie geen verschillen zijn tussen werknemers en zelfstandigen, is wel te zien dat passieve en actieve politieke participatie op verschillende andere manieren worden beïnvloed. Aangevoerd is namelijk dat hoog opleidingsniveau van grotere invloed is op actieve politieke participatie dan op passieve politieke participatie. Daarnaast is te zien dat de invloed van het hebben van een migratieachtergrond verschilt: mensen met een migrantenachtergrond hebben ten opzichte van mensen zonder migratieachtergrond een lagere kans op passieve politieke participatie, terwijl er geen verschillen zijn in de kans op actieve politieke participatie. Het is dus belangrijk om onderscheid te maken tussen actieve en passieve politieke participatie, want deze worden op verschillende manieren beïnvloed.

Op deze manier kunnen we terugkoppelen naar de hoofdvraag: zelfstandig ondernemerschap en werknemerschap zorgen niet voor verschillen in politieke participatie, ook niet in combinatie met opleidingsniveau. Daardoor is er ook geen indicatie dat er verschillen zijn in de mate waarin goalframes een rol spelen politieke participatie te verklaren. Dat is het geval

bij zowel actieve als passieve vormen van politieke participatie. Wel zijn er verschillen in de manieren waarop actieve en passieve politieke participatie verklaard worden. Er is dus reden om actieve en passieve participatie als twee verschillende concepten te behandelen.

5.2 Discussie

Het onderzoek kent wel een aantal beperkingen. Allereerst zijn er wat beperkingen bij de operationalisaties van de variabelen. Zo is passieve politieke participatie minder breed gemeten dan dat het is gedefinieerd. In dit onderzoek is namelijk alleen het stemmen tijdens de laatste Tweede Kamerverkiezingen als meting voor passieve politieke participatie genomen. Echter, er zijn meerdere vormen van passieve politieke participatie niet gemeten, zoals het tekenen van een petitie. Toch is stemmen wel de meest gangbare vorm van passieve politieke participatie, waardoor het een redelijke maatstaf is voor het concept. Een ander punt van kritiek geldt voor beide vormen van politieke participatie: er is alleen gemeten of mensen hebben geparticipeerd en niet hoe vaak of op hoeveel verschillende manieren ze participeren. Het gaat daarom om de kans op politieke participatie, niet zo zeer de frequentie daarvan. De verschillen in politieke participatie zouden daardoor kunnen worden onderschat. Een andere beperking is dat het individu misschien niet in staat was om tijdens de laatste verkiezingen te gaan stemmen vanwege ziekte of vakantie, dus dat mensen deze ene mogelijkheid op passieve politieke participatie niet konden pakken. Hierdoor zou de groep die niet heeft gestemd overschat worden. Echter, we verwachten dat deze groep redelijk klein is en bovendien gelijk voor alle groepen in de populatie, waardoor het niet aannemelijk is dat ze van invloed zijn op de uitkomsten van het onderzoek. De meting van zelfstandige/werknemer is ook een beperking van dit onderzoek, want als werksituatie is genomen welke situatie op dit moment het meest van toepassing is. Daarmee zou een respondent pas net zelfstandige geworden zijn of zelfstandig ondernemerschap met werknemerschap kunnen combineren. Op die manier zijn de verschillen tussen een zelfstandige en een werknemer minder groot.

Het ontbreken van een volledige onafhankelijke steekproeftrekking was ook een beperking van het onderzoek. Er kunnen namelijk respondenten uit dezelfde huishoudens in de steekproef gekomen zijn, waardoor die elkaars resultaten kunnen beïnvloeden. Maar, deze beperking hebben we ondervangen door het significantieniveau te verhogen, waarmee de overschatte steekproefomvang is tegengegaan. Tot slot is een beperking van het onderzoek dat de modellen geen goede voorspellers van de kansen op politieke participatie zijn.

Desalniettemin wil dat niet zeggen dat de gevonden verbanden niet bestaan, maar alleen dat er nog veel ruis bestaat. Er kunnen dus nog andere verbanden bestaan die van invloed zijn op de

politieke participatie die niet in deze modellen zijn opgenomen. Te denken valt aan verschillen in politiek vertrouwen, geïnformeerde status of sociaal economische status. Deze onzichtbare verbanden zouden van invloed kunnen zijn op de verbanden in de modellen. Er is dus nog enige voorzichtigheid geboden om de invloed van de goalframes in het geheel uit te sluiten.

Ondanks deze beperkingen kunnen we met de resultaten redelijke uitspraken doen over de theorie. Dat komt ook doordat de resultaten goed generaliseerbaar zijn naar de huidige Nederlandse bevolking. De dataset bevat namelijk een representatieve steekproef van de Nederlandse bevolking uit 2022 en verdere selectie voor dit onderzoek heeft niet geleid tot sterke veranderingen in de verdeling van deze steekproef.

Dat de theorie niet ondersteund wordt door de resultaten, kunnen we op verschillende manieren verklaren. Allereerst is de theorie gestoeld op de aanname dat er verschillen zijn in de dominantie van goalframes van werknemers en zelfstandigen. Wanneer die aanname niet klopt, dan zouden we volgens de rest van de theorie ook geen verschillen in politieke participatie vinden. Het is dus mogelijk dat we de verschillen in de werksituatie van zelfstandigen en werknemers hebben overschat. Zo was het namelijk mogelijk dat iemand in de steekproef zowel werknemer als zelfstandige was, waardoor de onderlinge verschillen kleiner zijn geworden. Daarnaast is er een theoretisch argument dat de verschillen niet zo klein zijn als we dachten, doordat steeds meer mensen er voor kiezen om zich als zelfstandige te verhuren voor werk dat normaal door werknemers wordt gedaan (Kösters, 2023). Daarmee kunnen ze alsnog werken met ‘collega’s’ ook al hebben ze geen arbeidscontract. De verschillen tussen werknemers en zelfstandigen zijn op die manier verkleind, waardoor de invloed van de *normative goalframe* is verkleind. Een andere mogelijke reden voor dat we de verwachte verbanden niet hebben kunnen aantonen is dat er sprake is van andere niet-zichtbare verbanden zijn die de verwachte verbanden hebben onderdrukt. Er is namelijk nog veel ruis, dus er kunnen nog andere factoren een rol spelen die we nu niet hebben waargenomen. Voorbeelden hiervan zijn een verschil in politiek vertrouwen, politieke geïnformeerde status of sociaal economische status. Wanneer daarin verschillen zijn tussen zelfstandigen en werknemers, dan zouden die van invloed kunnen zijn op onze resultaten, zonder dat we dat hebben waargenomen.

Het onderzoek biedt verschillende aanknopingspunten voor vervolgonderzoek. Allereerst zou *vervolgonderzoek* gericht kunnen kijken naar de invloed van de *normative* en de *gain* goalframe op politieke participatie. Een mogelijkheid zou zijn een kwalitatief onderzoek naar

de beweegredenen voor politieke participatie. De gegeven beweegredenen kunnen gekoppeld worden aan de manier waarop mensen de situatie interpreteren: als een sociale situatie of als een situatie met economische mogelijkheden dan wel bedreigingen. Dit kan dan gereflecteerd worden aan de goalframes. Een mogelijk probleem daarvan is wel dat mensen eerder sociaal gewenste antwoorden zouden geven. Een andere onderzoeksrichting zou zijn om de verschillen tussen passieve en actieve politieke participatie verder te onderzoeken, specifiek de relatie met migratieachtergrond. Hierbij kan onderzocht worden wat de reden is dat het hebben van een migratieachtergrond zorgt voor een lagere passieve politieke participatie, terwijl er geen verschillen zijn inactieve politieke participatie. Hebben immigranten misschien minder vertrouwen in de conventionele politiek, waardoor ze op zoek gaan naar alternatieven om de politiek te beïnvloeden? Tot slot is ook de verhouding tussen actieve en passieve politieke participatie een goede onderzoeksrichting. Men kan namelijk de vraag stellen of actieve politieke participatie een substituut is van passieve politieke participatie of een aanvulling daarop. Dus kan onderzocht worden of de groep die doet aan actieve politieke participatie dezelfde groep is die doet aan passieve politieke participatie of dat dat juist een andere groep is.

Dit vervolgonderzoek zou daarmee een bijdrage kunnen leveren aan eventueel beleid dat er gericht op zou zijn om politieke participatie bij groepen met een lage participatie te stimuleren. Wanneer we inzicht hebben verkregen in welke goalframe een grotere rol speelt bij de keuze van individuen om te gaan politiek participeren, kunnen we beleid meer gericht inzetten om in te spelen op de prikkels en doelen die bij die goalframe horen.

6. Bronvermelding

- Becker, R. (2004). Political efficacy and voter turnout in East and West Germany. *German Politics, 13*, 317-340. doi: 10.1080/0964400042000248223
- Boyd, R. W. (1989). The effects of primaries and statewide races on voter turnout. *The Journal of Politics, 51* (3), 730-739.
- Brady, H., Verba, S., & Schlozman, K. (1995). Beyond SES: a resource model of political participation. *American Political Science Review, 89* (2), 271-294. doi:10.2307/2082425
- Brennan, G. & Pettit, P. (2000). The hidden economy of esteem. *Economy and Philosophy, 16*, 77-98. doi:10.1017/S0266267100000146
- Brewer, G. A. (2003). Building social capital: civic attitudes and behavior of public servants. *Journal of Public Administration Research and Theory, 13* (1), 5-26.
<https://doi.org/10.1093/jpart/mug011>
- Centraal Bureau voor de Statistiek (2021)a. Wie zijn de zzp'ers?
- Centraal Bureau voor de Statistiek (2021)b. Wat voor werk doen zzp'ers?
- Centraal Bureau voor de Statistiek (2022). Arbeidsdeelname; migratieachtergrond, 2003-2022. <https://www.cbs.nl/nl-nl/cijfers/detail/82809NED>
- Dalton, R. (1984). Cognitive mobilisation and partisan dealignment in advanced industrial democracies. *The Journal of Politics, 46*, 264-284.
- Ertas, N. (2015). Political voice and civic attentiveness of public and non-profit employees. *American Review of public administration, 45* (5), 607-626.
<https://doi.org/10.1177/0275074014523117>
- Gilbert, C. P. (1993). *The impact of churches on political behavior: an empirical study*. Westport, CT: Greenwood Press.
- Hadjar, A., & Beck, M. (2010). Who does not participate in elections in Europe and why is this? *European Societies, 12*(4), 521-542. DOI: 10.1080/14616696.2010.483007
- Kenny, C. B. (1992). Political participations and effects from the social environment. *American Journal of Political Science, 36* (1), 259-267.

- Kösters, L. (2023). Aantal zelfstandigen zonder personeel onder 25 jaar sterk gegroeid. <https://esb.nu/aantal-zelfstandigen-zonder-personeel-onder-25-jaar-sterk-gegroeid/>
- Leighley, J. (1990). Social Interaction and contextual influences on political participation. *American Politics Quarterly*, 18, 459-475.
- Leighley, J. (1995). Attitudes, opportunities and incentives: A field essay on political participation. *Political Research Quarterly*, 48(1), 181-209. doi: 10.1177/106591299504800111
- Lind, G. (2006). Effective moral education: The Konstanz method of dilemma discussion. *Hellenic Journal of Psychology*, 3, 189-196.
- Lindenberg, S. & Steg, L. (2007). Normative, gain and hedonic goalframes guiding environmental behavior. *Journal of Social Issues*, 63(1), 117-137.
- Mencken, F.C., Smith, B. & Tolbert, C.M. (2020). Self-employment and civic inclination. *Sociological Perspectives*, 63(5), 719-737.
- Milinski, M. (2016). Reputation, a universal currency for human social interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371, 1687. <https://doi.org/10.1098/rstb.2015.0100>
- Milligan, K., Moretti, E. & Oreopoulos, P. (2004). Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics*, 88, 1667-1695. <https://doi.org/10.1016/j.jpubeco.2003.10.005>
- Nie, N.H., Junn J. & Stehlik-Barry, K. (1996). *Education and democratic citizenship in America*. Chicago. IL: University Press of Chicago.
- Oppenhuis, E. (1995). *Voting behavior in Europe: A comparative analysis of electoral participation and party choice*. Amsterdam, the Netherlands: Het Spinhuis.
- Ostrom, E. (2000). Collective action and the evolution of social Norms. *The Journal of Economic perspectives*, 14 (3), 137-158. <https://doi-org.proxy-ub.rug.nl/10.1257/jep.14.3.137>
- Putnam, R.D. (2000). *Bowling alone: the collapse and revival of American community*. New York: Simon and Schuster.

- Rotolo, T. & Wilson, J. (2006). Employment sector and volunteering: The contribution of nonprofit and public sector workers to the volunteer labor force. *Sociological Quarterly*, 47, 21-40. <https://doi-org.proxy-ub.rug.nl/10.1111/j.1533-8525.2006.00036.x>
- Rosenstone, S.J. & Hansen, J.M. (1993). Mobilization, participation and democracy in America. *American Political Science Review*, 88 (3), 771. doi:10.2307/2944841
- Sairambay, Y. (2020). Reconceptualising political participation. *Human Affairs*, 30(1), 120-127. <https://doi-org.proxy-ub.rug.nl/10.1515/humaff-2020-0011>
- Schmeets, H., Marcellino, C. & Conceiao J.F. (2021). De relevantie van verblijfsduur van immigranten voor hun participatie in de samenleving. CBS. <https://www.cbs.nl/nl-nl/longread/statistische-trends/2021/de-relevantie-van-verblijfsduur-van-immigranten-voor-hun-participatie-in-de-samenleving?onepage=true>
- Sociaal-Economische Raad (2010). Zzp'ers in beeld. Een integrale visie op zelfstandigen zonder personeel.
- Simpson, B. & Willer, R. (2015). Beyond altruism: sociological foundations of cooperation and prosocial behavior. *Annual Review of Sociology*, 41, 43-63. <http://proxy-ub.rug.nl/login?url=http://dx.doi.org/10.1146/annurev-soc-073014-112242>
- Van der Meer, T. & Dekker P. (2011). Politieke participatie. In: *Respect man! Tolerantie, solidariteit en andere modern woorden* (211-266). Celsus.
- Wilson, J., & Musick, M. A. (1997). Work and Volunteering: The Long Arm of the Job. *Social Forces*, 76(1), 251–272. <https://doi.org/10.2307/2580325>

Bijlage 1 Operationalisatie van de variabelen

In deze bijlage wordt uitgelegd op welke manier de gebruikte variabelen zijn geoperationaliseerd. Daarbij behandelen we a) de originele variabele en b) de bewerkingen en de uiteindelijke variabele. We geven de statistieken van de variabele van vóór de selectie, zodat in bijlage 2 de steekproef kan worden vergeleken met het geheel van de data. Tenslotte bespreken we de selectie: hoe is deze uitgevoerd, waarom en wat is het gevolg ervan voor de steekproefomvang. Voor elk van deze stappen zijn de *syntax* en de *output* weergegeven.

De gebruikte variabelen in dit onderzoek zijn:

1. Passieve politieke participatie
2. Actieve politieke participatie
3. Opleidingsniveau
4. Werksituatie
5. Migratieachtergrond
6. Geslacht
7. Leeftijd

1. Passieve politieke participatie

a. Originele variabele

Voor de afhankelijke variabele passieve politieke participatie is de variabele *cv22n053* gebruikt, daarbij is gevraagd: *Hebt u tijdens de laatste Tweede Kamerverkiezingen op 17 maart 2021 gestemd?* De antwoordcategorieën zijn: 1) ja, 2) nee, 3) ik was niet stemgerechtigd en -9) ik weet het niet. De variabele is als volgt verdeeld:

```
FREQUENCIES VARIABLES=cv22n053  
/ORDER=ANALYSIS.
```

cv22n053 Did you vote in the most recent parliamentary elections, held on 17 March 2021?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 yes	4424	38,0	86,6	86,6
	2 no	573	4,9	11,2	97,8
	3 not eligible to vote	113	1,0	2,2	100,0
	Total	5110	43,9	100,0	
Missing	-9 I dont know	57	,5		
	System	6475	55,6		
	Total	6532	56,1		
Total		11642	100,0		

b. Bewerkingen en uiteindelijke variabele

Om de variabele bruikbaar te maken voor het onderzoek is het gecodeerd naar 'Passieve' met de scores van 1) ja 0) nee. Een score van 1 betekent dus dat een respondent heeft gestemd en een score van 0 betekent dat een respondent niet heeft gestemd in 2017. Door de bewerking is de variabele dus gedraaid. De categorieën 3) ik was niet stemgerechtigd en -9) ik weet het niet zijn als *missing* gecodeerd. De verdeling tussen wel- en niet-stemmers is dus niet veranderd, zoals ook is te zien in de frequentieverdeling hieronder.

RECODE cv22n053 (2=0) (1=1) (ELSE=sysmys) INTO Passieve.
EXECUTE.

FREQUENCIES VARIABLES=passieve
/ORDER=ANALYSIS.

Passieve					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	573	4,9	11,5	11,5
	1,00	4424	38,0	88,5	100,0
	Total	4997	42,9	100,0	
Missing	System	6645	57,1		
Total		11642	100,0		

2. Actieve politieke participatie

a. Originele variabele

De tweede afhankelijke variabele is geconstrueerd uit zes verschillende items die voortkomen uit de vraag: *Er zijn verschillende manieren om iets politiek aan de orde te stellen of invloed te hebben op politici of de overheid. Wilt u van de volgende mogelijkheden aangeven of u deze in de afgelopen 5 jaar hebt gebruikt? (Meerdere antwoorden mogelijk)*. De zeven items zijn: cv22n065 hulp gevraagd aan radio, televisie of krant; cv22n066 een politieke partij of organisatie ingeschakeld, cv22n067 meegedaan aan een door de overheid georganiseerde

inspraakbijeenkomst, hoorzitting of discussiebijeenkomst; *cv22n068* contact opgenomen met een politicus of ambtenaar; *cv22n069* meegedaan aan een actiegroep; *cv22n070* meegedaan aan een protestactie, protestmars of demonstratie en *cv22n071* via internet, e-mail of sms meegedaan aan een politieke discussie of actie. Per item hadden de respondenten de volgende antwoordmogelijkheden: 0) nee, 1) ja en -9) ik weet het niet. De items hebben de volgende verdelingen:

FREQUENCIES VARIABLES= *cv22n065 cv22n066 cv22n067 cv22n068 cv22n069 cv22n070 cv22n071*
 /ORDER=ANALYSIS.

cv22n065 Raising a political issue or influencing politicians or government: asked for help from radio, television or newspaper

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 no	4937	42,4	97,0	97,0
	1 yes	153	1,3	3,0	100,0
	Total	5090	43,7	100,0	
Missing	-9 I dont know	188	1,6		
	System	6364	54,7		
	Total	6552	56,3		
Total		11642	100,0		

cv22n066 Raising a political issue or influencing politicians or government: by making use of a political party or organization

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 no	4987	42,8	98,0	98,0
	1 yes	103	,9	2,0	100,0
	Total	5090	43,7	100,0	
Missing	-9 I dont know	188	1,6		
	System	6364	54,7		
	Total	6552	56,3		
Total		11642	100,0		

cv22n067 Raising a political issue or influencing politicians or government: participated in a government-organized public hearing, discussion or citizens participation meeting

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 no	4914	42,2	96,5	96,5
	1 yes	176	1,5	3,5	100,0
	Total	5090	43,7	100,0	
Missing	-9 I dont know	188	1,6		
	System	6364	54,7		
	Total	6552	56,3		
Total		11642	100,0		

cv22n068 Raising a political issue or influencing politicians or government: contacted a politician or civil servant

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 no	4879	41,9	95,9	95,9
	1 yes	211	1,8	4,1	100,0
	Total	5090	43,7	100,0	
Missing	-9 I dont know	188	1,6		
	System	6364	54,7		
	Total	6552	56,3		
Total		11642	100,0		

cv22n069 Raising a political issue or influencing politicians or government: participated in an action group

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 no	4984	42,8	97,9	97,9
	1 yes	106	,9	2,1	100,0
	Total	5090	43,7	100,0	
Missing	-9 I dont know	188	1,6		
	System	6364	54,7		
	Total	6552	56,3		
Total		11642	100,0		

cv22n070 Raising a political issue or influencing politicians or government: participated in a protest action, protest march or demonstration

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 no	4914	42,2	96,5	96,5
	1 yes	176	1,5	3,5	100,0
	Total	5090	43,7	100,0	
Missing	-9 I dont know	188	1,6		
	System	6364	54,7		
	Total	6552	56,3		
Total		11642	100,0		

cv22n071 Raising a political issue or influencing politicians or government: participated in a political discussion or campaign on the Internet, by e-mail or SMS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 no	4568	39,2	89,7	89,7
	1 yes	522	4,5	10,3	100,0
	Total	5090	43,7	100,0	
Missing	-9 I dont know	188	1,6		
	System	6364	54,7		
	Total	6552	56,3		
Total		11642	100,0		

b. Bewerkingen en uiteindelijke variabele

Voor elk item is eerst -9) ik weet het niet als *missing* gecodeerd. Vervolgens zijn de items bij elkaar opgeteld tot een somscore. Deze somscore is vervolgens gedummificeerd, zodat iedere score hoger dan 0 gelijk is aan 1. Op deze manier heeft de uiteindelijke variabele twee categorieën, waarbij 1) staat voor meegedaan aan een mogelijkheid voor actieve politieke participatie en 0) niet meegedaan aan een mogelijkheid. Mensen die vaker aan een of meer verschillende mogelijkheden hebben meegedaan krijgen dus ook een waarde van 1. Te zien aan de uiteindelijke variabele is dat de meeste mensen niet meedoen aan vormen van actieve politieke participatie. Dat komt overeen met de items.

```
RECODE cv22n065 cv22n066 cv22n067 cv22n068 cv22n069 cv22n070 cv22n071 (-9=SYSMIS)
(ELSE=Copy) INTO Media partij bijeenkomst politicus actiegroep protest online.
EXECUTE.
```

```
COMPUTE Actiefvoor=Media+partij+bijeenkomst+politicus+actiegroep+protest+online.
EXECUTE.
```

RECODE Actiefvoor (0=0) (1 thru 7=1) (ELSE=SYSMIS) INTO Actieve.
EXECUTE.

FREQUENCIES VARIABLES=Actieve
/ORDER=ANALYSIS.

		Actieve			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	4159	35,7	81,7	81,7
	1,00	931	8,0	18,3	100,0
	Total	5090	43,7	100,0	
Missing	System	6552	56,3		
Total		11642	100,0		

3. Werksituatie (zelfstandige en werknemer private sector)

a. Originele variabele (zelfstandige)

Werksituatie bestaat uit twee dummy's. Deze zijn op verschillende manieren geconstrueerd. Eerst behandelen we hier zelfstandige.

Voor de variabele zelfstandige zijn twee variabelen gebruikt. Bij de eerste variabele, *cw21n121*, is gevraagd: *Bent u werknemer in vaste of tijdelijke dienst, oproepkracht, uitzendkracht of bent u een zelfstandige/freelancer of vrije beroepsbeoefenaar? Het gaat hier om de baan waarin u de meeste uren werkt, (andere) (bij-)banen komen later aan de orde. Had u meerdere (hoofd)banen die wat uren betreft gelijk zijn, neem dan de baan die voor u het belangrijkste is. Een directeur van een NV of BV is meestal in loondienst bij de NV of BV. Vul dan in dat u werknemer in vaste of tijdelijke dienst bent. Ook een directeur-groootaandeelhouder ontvangt gewoonlijk een inkomen als werknemer. Toch willen we u – als dat van toepassing is – vragen aan te geven dat u directeur(-groootaandeelhouder) bent.* De antwoordopties zijn: 1) werknemer in vaste dienst, 2) werknemer in tijdelijke dienst, 3) oproepkracht, 4) uitzendkracht, 5) zelfstandige/freelancer, 6) vrije beroepsbeoefenaar, 7) directeur NV of BV, 8) directeur-groootaandeelhouder. De andere variabele is een variabele die voortkomt uit twee vervolgvragen.

De respondenten die geantwoord hadden dat ze 5) zelfstandige/freelancer of 6) vrije beroepsbeoefenaar zijn, kregen de vervolgvraag *cw21n124*: *Participeert/participeerde u in een maatschap of vennootschap onder firma (VOF)? Huisartsen, notarissen en agrariërs kennen meestal een maatschapvorm; productiebedrijven meestal een vennootschap onder firma (VOF).* De antwoordcategorieën zijn 1) ja, in een maatschap, 2) ja, in een VOF en 3)

nee. Wanneer ze daar 3) nee als antwoord hadden gegeven, kregen ze de vraag *cw21n125: Bent/was u zelfstandige of werkt/werkte u mee in het gezinsbedrijf of was u freelancer?* De antwoordopties zijn 1) zelfstandige, 2) werk(te) mee in gezinsbedrijf en 3) freelancer. Deze variabelen zagen er als volgt uit.

FREQUENCIES VARIABLES=cw21n121 cw21n125
/ORDER=ANALYSIS.

cw21n121 Are (/In your last job) (/were) you an employee in permanent or temporary employment, an on-call employee or temp-staffer, or are (/were) you a self-employed/freelance worker, or an independent professional?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 employee in permanent employment	2560	22,0	72,8	72,8
	2 employee in temporary employment	418	3,6	11,9	84,7
	3 on-call employee	99	,9	2,8	87,5
	4 temp-staffer	75	,6	2,1	89,6
	5 self-employed/freelancer	292	2,5	8,3	97,9
	6 independent professional	28	,2	,8	98,7
	7 director of a limited liability or private limited company	9	,1	,3	98,9
	8 majority shareholder director	37	,3	1,1	100,0
	Total	3518	30,2	100,0	
Missing	System	8124	69,8		
Total		11642	100,0		

cw21n125 Are (were) you self-employed or do (did) you work in a family business, or are (were) you a freelancer?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 self-employed	200	1,7	75,2	75,2
	2 work(ed) in a family business	5	,0	1,9	77,1
	3 freelancer	61	,5	22,9	100,0
	Total	266	2,3	100,0	
Missing	System	11376	97,7		
Total		11642	100,0		

b. Bewerking en uiteindelijke variabele

Om de variabele zelfstandige te creëren zijn de twee variabelen samengevoegd. Maar eerst is de eerste variabele (*cw21n121*) gecodeerd tot 0) werknemer en 1) zelfstandige en -5) anders. Onder 0) werknemer zijn de originele categorieën 1) werknemer in vaste dienst, 2) werknemer in tijdelijke dienst en 3) oproepkracht opgenomen, omdat dit mensen zijn met een vaste werkgever. 4) Uitzendkrachten zijn niet meegenomen omdat ze geen vaste werkgever hebben. Onder 1) zelfstandige vallen de originele categorieën 5) zelfstandige/freelancer en 6) vrije beroepsbeoefenaar. De andere categorieën vallen onder -5) anders. Vervolgens is de variabele *cw21n125* gecodeerd naar 0) zelfstandige/freelancer en -5) werkte mee in een gezinsbedrijf en de *missings*. Hierna zijn de scores van beide variabelen opgeteld in een nieuwe variabele: zelfstandige met als categorieën 1) zelfstandige, 0) werknemer en waarden lager dan 0 als 'anders'. Tot slot coderen we de categorie 'anders' als *missing*. Zo heeft de uiteindelijke variabele 3077 werknemers en 261 zelfstandigen.

```
RECODE cw21n121 (4=-5) (1 thru 3=0) (5 thru 6=-1) (7 thru 8=-5) (ELSE=Copy) INTO zelf1.
EXECUTE.
```

```
RECODE cw21n125 (1=2) (2=-5)(3=2)(else=0) INTO zelf2.
EXECUTE.
```

```
COMPUTE Zelf3=zelf1 + zelf2.
EXECUTE.
```

```
RECODE Zelf3 (-10 thru -1=SYSMIS) (ELSE=Copy) INTO Zelfstandige.
EXECUTE.
```

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

		Zelfstandige			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	3077	26,4	92,2	92,2
	1,00	261	2,2	7,8	100,0
	Total	3338	28,7	100,0	
Missing	System	8304	71,3		
Total		11642	100,0		

- a. Originele variabele (werknemer private sector)

Voor werknemer uit de private sector is de volgende variabele gebruikt: *cw21n122*. Voor deze variabele werd gevraagd: *In welk type organisatie werkt u?* De antwoordopties zijn 1) overheid, semi-overheid en 2) particulier bedrijf. Deze variabele ziet er zo uit:

FREQUENCIES VARIABLES=*cw21n122*
/ORDER=ANALYSIS.

In what type of organization do (did) you work (in your last job)?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	public, semi-public sector	914	35,1	38,0	38,0
	private company	1491	57,3	62,0	100,0
	Total	2405	92,4	100,0	
Missing	System	199	7,6		
Total		2604	100,0		

b. Bewerkingen en de uiteindelijke variabele

Om de uiteindelijke variabele te construeren is *cw21n122* gecodeerd naar 1) private sector en 0) anders. Belangrijk is dat ook de *missings* als 0) anders zijn gecodeerd, omdat anders zelfstandigen allemaal als *missing* zijn gecodeerd en deze variabele daarmee onbruikbaar zou zijn voor de uiteindelijke modellen. De uiteindelijke variabele ziet er dan zo uit:

recode *cw21n122* (1=0)(2=1)(else=0) into *wnpriv*.
execute.

FREQUENCIES VARIABLES=*wnpriv*
/ORDER=ANALYSIS.

		wnpriv			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	9642	82,8	82,8	82,8
	1,00	2000	17,2	17,2	100,0
Total		11642	100,0	100,0	

4. Opleidingsniveau

a. Originele variabele

Voor opleidingsniveau is de achtergrond variabele *oplmet* gebruikt. Deze variabele beschrijft wat de hoogst genoten opleiding met diploma van de respondent is. De categorieën zijn: 1) basisonderwijs, 2) vmbo, 3) havo/vwo, 4) mbo, 5) hbo, 6) wo, 7) anders, 8) (nog) geen onderwijs afgerond en 9) volgt nog geen onderwijs.

FREQUENCIES VARIABLES=oplmet
/ORDER=ANALYSIS.

oplmet Highest level of education with diploma

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 primary school	685	5,9	6,0	6,0
	2 vmbo (intermediate secondary education, US: junior high school)	1785	15,3	15,7	21,7
	3 havo/vwo (higher secondary education/preparatory university education, US: senior high school)	1011	8,7	8,9	30,6
	4 mbo (intermediate vocational education, US: junior college)	2362	20,3	20,8	51,4
	5 hbo (higher vocational education, US: college)	2470	21,2	21,7	73,2
	6 wo (university)	1330	11,4	11,7	84,9
	7 other	188	1,6	1,7	86,6
	8 Not (yet) completed any education	894	7,7	7,9	94,4
	9 Not yet started any education*	633	5,4	5,6	100,0
	Total	11358	97,6	100,0	
Missing	System	284	2,4		
Total		11642	100,0		

b. Bewerking en uiteindelijke variabele

Omdat een aantal categorieën in de uiteindelijke steekproef te klein worden, is ervoor gekozen om categorieën samen te voegen en twee dummy variabelen te maken. Allereerst is de dummy *OPLlaag* gemaakt, waarbij de categorieën van 1, 2, 8 en 9 als 1 zijn gecodeerd en de anderen als 0. Daarna is de dummy *OPLhoog* gemaakt, waarbij de categorieën 5 en 6 zijn gecodeerd als 1 en de anderen als 0. Hierdoor wordt een gemiddeld opleidingsniveau de referentiecategorie, bestaande uit 3 en 4. Uiteindelijk zijn er 3997 respondenten als laagopgeleid geclassificeerd, 3800 als hoogopgeleid en 3373 als gemiddeld opgeleid.

Recode oplmet (1 thru 2 =1) (8 thru 9=1)(3 thru 6=0) (7=systemis) into OPLlaag.
execute.

recode oplmet (5 thru 6 =1)(1 thru 4=0) (8 thru 9=0) (7=sysmis) into OPLhoog.
execute.

FREQUENCIES VARIABLES= OPLlaag OPLhoog
/ORDER=ANALYSIS.

OPLlaag

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	7173	61,6	64,2	64,2
	1,00	3997	34,3	35,8	100,0
	Total	11170	95,9	100,0	
Missing	System	472	4,1		
Total		11642	100,0		

OPLhoog

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	7370	63,3	66,0	66,0
	1,00	3800	32,6	34,0	100,0
	Total	11170	95,9	100,0	
Missing	System	472	4,1		
Total		11642	100,0		

5. Leeftijd

a. Originele variabele

Voor de controlevariabele leeftijd is de achtergrondvariabele leeftijd gebruikt. Deze variabele is als volgt verdeeld.

FREQUENCIES VARIABLES=leeftijd
/FORMAT=NOTABLE
/NTILES=4
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
/ORDER=ANALYSIS.

Statistics

leeftijd Age of the household membe		
N	Valid	11358
	Missing	284
Mean		44,30
Std. Deviation		22,948
Minimum		0
Maximum		122
Percentiles	25	25,00
	50	46,00
	75	64,00

b. Bewerking en de uiteindelijke variabele.

Allereerst zijn leeftijden onder de 18 jaar en 110 jaar of ouder als *missings* gecodeerd. Respondenten van jonger dan 18 jaar mogen namelijk nog niet stemmen en kunnen daardoor de uitkomsten van ons onderzoek vertekenen. Waarden hoger dan 110 beschouwen we als niet-realistische waarden. Bovendien is het onrealistisch dat deze groepen werken. Hiermee stijgt het aantal *missings* met 1.821, waarvan 1.819 onder de 18 en 2 boven de 110. Vervolgens is de variabele gecentreerd op het gemiddelde, zodat in de regressiemodellen de coëfficiënt van leeftijd meer waardevol is. Dat is zo omdat nu elke stap van leeftijd waardevol is, in plaats van dat er eerst 18 stappen moeten worden genomen voordat de waarde relevant is.

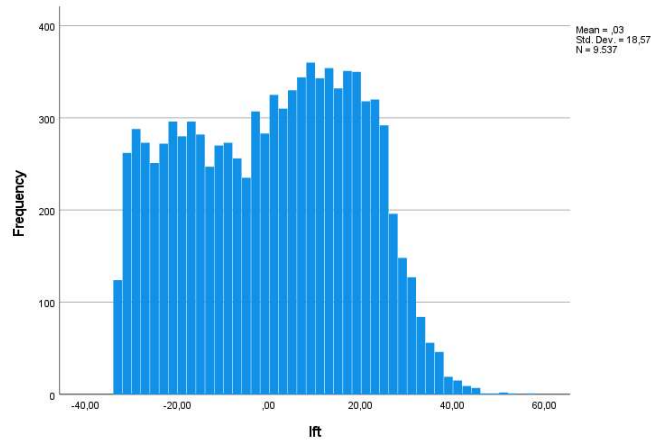
recode leeftijd (0 thru 17=sysmis)(110 thru 150=sysmis)(else=copy) into leef.
execute.

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

Compute lft= leef-50.9.
execute.

```
FREQUENCIES VARIABLES=lft
/FORMAT=NOTABLE
/NTILES=4
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
/ORDER=ANALYSIS.
```


Statistics		
lft		
N	Valid	9537
	Missing	2105
Mean		,0322
Std. Deviation		18,56972
Minimum		-32,90
Maximum		57,10
Percentiles	25	-15,9000
	50	1,1000
	75	15,1000



6. Geslacht

a. Originele variabele

Voor de variabele geslacht is de achtergrond variabele daarvan gebruikt. De antwoordcategorieën zijn 1) man, 2) vrouw en 3) anders.

FREQUENCIES VARIABLES= geslacht
/ORDER=ANALYSIS.

geslacht Gender					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	5574	47,9	49,1	49,1
	2 Female	5771	49,6	50,8	99,9
	3 Other	13	,1	,1	100,0
	Total	11358	97,6	100,0	
Missing	System	284	2,4		
Total		11642	100,0		

b. Bewerkingen en uiteindelijke variabele

De variabele is bewerkt zodat de uiteindelijke categorieën zijn 0) vrouw en 1) man, waarbij anders als *missing* is gecodeerd. Deze heeft namelijk te weinig respondenten om relevante uitspraken over te kunnen doen. Om de interpretatie van de coëfficiënt van de variabele in de modellen makkelijker te maken is de variabele *man* genoemd.

recode geslacht (2=0)(1=1) (3=sysmis) (else=copy) into man.
execute.

FREQUENCIES VARIABLES= man
/ORDER=ANALYSIS.

		man			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	5771	49,6	50,9	50,9
	1,00	5574	47,9	49,1	100,0
	Total	11345	97,4	100,0	
Missing	System	297	2,6		
Total		11642	100,0		

7. Migratieachtergrond

a. Originele variabele

Migratieachtergrond is geconstrueerd uit de achtergrondvariabele *herkomstgroep*. Deze variabele heeft de volgende categorieën: 0) autochtone herkomst, 101) eerste generatie van niet-Nederlandse, westerse afkomst, 102) eerste generatie van niet-Nederlandse, niet-westerse afkomst, 201) tweede generatie van niet-Nederlandse, westerse afkomst en 202) tweede generatie van niet-Nederlandse, niet-westerse afkomst.

FREQUENCIES VARIABLES= herkomstgroep
/ORDER=ANALYSIS.

		herkomstgroep Origin			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 Dutch background	6366	54,7	81,5	81,5
	101 First generation foreign, Western background	298	2,6	3,8	85,3
	102 First generation foreign, non-western background	394	3,4	5,0	90,4
	201 Second generation foreign, Western background	427	3,7	5,5	95,9
	202 Second generation foreign, non-western background	323	2,8	4,1	100,0
	Total	7808	67,1	100,0	
Missing	System	3834	32,9		
Total		11642	100,0		

Om te voorkomen dat in de uiteindelijke analyse de groepen te klein zijn om zinvolle uitspraken over te kunnen doen, voegen we de groepen samen tot: geen migratieachtergrond, eerste generatie immigrant, tweede generatie immigrant. Om dat voor elkaar te krijgen, zijn er twee dummyvariabelen gemaakt: GEN1 met 1) eerste generatie en 0) anders, en GEN2 met 1)

tweede generatie en 0)anders. Op die manier is de groep zonder migratieachtergrond de referentiegroep. De variabelen zien er als volgt uit:

recode herkomstgroep (101 thru 102 =1) (0=0) (201 thru 202=0) (else=sysmis) into Gen1.
execute.

recode herkomstgroep (201 thru 202=1) (0 thru 102=0) (else=sysmis) into Gen2.
execute.

FREQUENCIES VARIABLES= gen1 gen2
/ORDER=ANALYSIS.

Gen1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	7116	61,1	91,1	91,1
	1,00	692	5,9	8,9	100,0
	Total	7808	67,1	100,0	
Missing	System	3834	32,9		
Total		11642	100,0		

Gen2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	7058	60,6	90,4	90,4
	1,00	750	6,4	9,6	100,0
	Total	7808	67,1	100,0	
Missing	System	3834	32,9		
Total		11642	100,0		

Selectie

De originele data heeft een steekproefomvang van 11.642 respondenten. Echter, niet alle data is bruikbaar voor ons onderzoek. Dat komt voornamelijk doordat door het samenvoegen van de databestanden er veel *missings* zijn. Om ervoor te zorgen dat we de geschatte regressiemodellen goed met elkaar kunnen vergelijken, is het nodig dat we voor de modellen dezelfde steekproef gebruiken. Om de selectie te maken schatten we een regressiemodel met alle variabelen, waarbij we een variabele opslaan met de *Cook's distance* waarde van elke respondent. Vervolgens gebruikten we die variabele als selectievariabele; dit betekent dat alleen respondenten met een score op deze variabele in de steekproef zijn gekomen. De steekproefomvang daalt hierdoor naar 2604.

REGRESSION
/DESCRIPTIVES MEAN STDDEV CORR SIG N
/MISSING LISTWISE

```
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT Zelfstandige
/METHOD=ENTER Actieve Passieve opleiding lft man nietnl
/SAVE COOK.
```

```
USE ALL.
FILTER BY COO_1.
EXECUTE.
```

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

Zelfstandige

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	2405	92,4	92,4	92,4
	1,00	199	7,6	7,6	100,0
Total		2604	100,0	100,0	

Naast de originele selectie gebruiken we voor modellen C en D, nog een extra selectie, namelijk op respondenten met een hoog opleidingsniveau. Na deze selectie is de steekproefomvang nog 1265. Deze selectie is als volgt gecodeerd:

```
COMPUTE filter_$=(OPLhoog = 1 & Coo_1>0).
VARIABLE LABELS filter_$ 'OPLhoog = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

Zelfstandige

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1150	90,9	90,9	90,9
	1,00	115	9,1	9,1	100,0
Total		1265	100,0	100,0	

Bijlage 2: Analyses

In deze bijlage tonen we alle analyses. Voor elke analyse geven we de *syntax* en de *output*, plus geven we een uitgebreide toelichting. Allereerst omschrijven we de univariate statistieken, daarbij beschrijven we de verdelingen en vergelijken we die met de verdelingen van vóór de selectie. Daarmee kunnen we beoordelen of de steekproef representatief is voor de dataset.

Vervolgens kijken we naar de bivariate statistieken om onderlinge verbanden tussen de variabelen te verkennen. Tenslotte schatten en bespreken we de binaire logistische modellen die we gebruiken. In bijlage 3 gaan we de kwaliteit van die modellen evalueren.

Univariate statistieken

```
FREQUENCIES VARIABLES=Zelfstandige Passieve Actieve man gen1 gen2 wnpriv OPLlaag
OPLhoog
/ORDER=ANALYSIS.
```

```
FREQUENCIES VARIABLES=ift
/NTILES=4
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
/HISTOGRAM
/ORDER=ANALYSIS.
```

Politieke participatie

		Passieve			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	268	10,3	10,3	10,3
	1,00	2336	89,7	89,7	100,0
Total		2604	100,0	100,0	

Het merendeel van de respondenten heeft wel gestemd tijdens de laatste Tweede Kamerverkiezingen in 2021: 89,7% tegenover 10,3%. Dit is ongeveer gelijk aan de originele verdeling waarbij 88,5% procent had gestemd tegenover 11,5% niet.

		Actieve			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	2141	82,2	82,2	82,2
	1,00	463	17,8	17,8	100,0
Total		2604	100,0	100,0	

17,9% van de respondenten in de steekproef heeft meegedaan aan vormen van actieve politieke participatie. Dat is ongeveer gelijk aan de 18,3% van de gehele dataset.

Werksituatie

Zelfstandige

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	2405	92,4	92,4	92,4
	1,00	199	7,6	7,6	100,0
Total		2604	100,0	100,0	

wnpriv

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1113	42,7	42,7	42,7
	1,00	1491	57,3	57,3	100,0
Total		2604	100,0	100,0	

Van de 2604 respondenten zijn er 199 zelfstandigen, wat neer komt op 7,6% van de respondenten. In de dataset was dat percentage ongeveer gelijk. Daarnaast werkt 57,3% als werknemer in de private sector en 35,1% als werknemer voor de overheid of een semi-overheid.

Opleidingsniveau

OPLlaag

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	2238	85,9	85,9	85,9
	1,00	366	14,1	14,1	100,0
Total		2604	100,0	100,0	

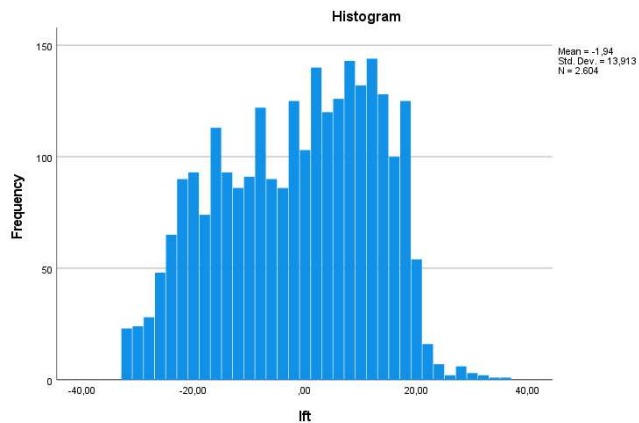
OPLhoog

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1339	51,4	51,4	51,4
	1,00	1265	48,6	48,6	100,0
Total		2604	100,0	100,0	

In de steekproef vallen 366 respondenten (14,1%) in de categorie van laag opleidingsniveau, 973 (37,3%) gemiddeld en 1265 (48,6%) hoog. Deze verdeling is anders dan bij de originele data, waar 34,1% een lage opleiding had, 33,3% een gemiddelde en 32,6% een hoge. Er zijn dus relatief veel laagopgeleiden afgevallen. Dat verklaren we doordat jongeren zonder afgeronde opleiding ook onder laagopgeleid vallen en niet meegenomen zijn in de analyses, omdat ze niet als werknemer of zelfstandige aan het werk zijn.

Leeftijd

Statistics		
lft		
N	Valid	2604
	Missing	0
Mean		-1,9357
Std. Deviation		13,91254
Minimum		-31,90
Maximum		35,10
Percentiles	25	-13,6500
	50	,1000
	75	10,1000



Van leeftijd, gecentreerd op 50,9, zien we dat de gemiddelde leeftijd -1,9 (daadwerkelijke leeftijd is 49) met een standaarddeviatie van 13,9. De variabele loopt van -31,9 tot 35,1, of zonder centrering van 19 tot 86. Bij de originele data is deze verdeling anders dan bij de originele steekproef waarbij de range loopt van 19 tot 86. Het gemiddelde is dan 0,0 met een standaarddeviatie 18,6. Dat deze standaarddeviatie veel groter is, is logisch en valt te verklaren doordat er alleen werkenden in de steekproef zijn meegenomen. Gepensioneerden en jongeren die nog met een opleiding bezig zijn, vallen daardoor af.

Geslacht

man					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1395	53,6	53,6	53,6
	1,00	1209	46,4	46,4	100,0
Total		2604	100,0	100,0	

In de steekproef is 46,4% man tegenover 53,6% vrouw. In de data is dit niet veel anders: 49,1% tegenover 50,1%.

Migratieachtergrond

		Gen1			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	2422	93,0	93,0	93,0
	1,00	182	7,0	7,0	100,0
Total		2604	100,0	100,0	

		Gen2			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	2337	89,7	89,7	89,7
	1,00	267	10,3	10,3	100,0
Total		2604	100,0	100,0	

Het grootste deel van de steekproef bestaat uit mensen zonder migratieachtergrond 82,7% tegenover 10,3% tweede generatie migranten en 7% eerste generatie. Deze percentages zijn een klein beetje verandert ten opzichte van de originele data: 81,5% zonder migratieachtergrond, 8,9% eerste generatie en 9,6% tweede generatie.

Conclusie

We kunnen dus concluderen dat de verdeling van de steekproef redelijk gelijk is, uitgezonderd de verklaarbare verschillen in leeftijd en opleidingsniveau. Gezien de originele steekproef representatief is voor de Nederlandse samenleving, kunnen we stellen dat deze uiteindelijke steekproef ook redelijk representatief is voor de Nederlandse samenleving.

Bivariate statistieken

Voor de bivariate statistieken berekenen we de bivariate correlaties van de variabelen. Vóór het berekenen van de statistieken coderen we de dummy's van 1) opleidingsniveau, 2) werksituatie en 3) migratieachtergrond als categoriale variabele waarbij de verdeling gelijk blijft. Op die manier zegt de associatiemaat meer over de verschillen tussen categorieën. De *syntaxen* van die bewerkingen zijn als volgt:

Opleidingsniveau: 0=laag, 1=gemiddeld, 2=hoog.

recode oplmet (3 thru 4=1)(1 thru 2=0) (8 thru 9 =0) (5 thru 6=2) (7=sysmis) into OPL .
execute.

Werksituatie 0=werknemer overheidssector, 1=werknemer private sector, 2=zelfstandige.

recode zelfstandige (1=2)(0=0) into zelf5.
execute.

Compute werksituatie= zelf5+wnpriv.

Migratieachtergrond: 0=geen migratieachtergrond, 1=eerste generatie, 2=tweede generatie

recode gen2 (1=2)(0=0) into mig.
execute.

compute migratieachtergrond = mig+gen1.
execute.

Vervolgens gebruiken we drie verschillende maten voor associatie: Cramer's V voor relaties tussen dummy's met categoriale variabelen of andere dummy's, F-toets voor gemiddelden voor relaties tussen categoriale variabelen en de continue variabele leeftijd en Pearson correlatie voor relaties tussen de continue variabele leeftijd en dummy's.

Per variabele gaan we nu de samenhang met de andere variabelen behandelen. Allereerst kijken we naar de samenhang van de afhankelijke variabelen onderling en de afhankelijke variabelen met de onafhankelijke variabelen. Vervolgens kijken we naar de samenhang van de onafhankelijke variabelen onderling. Mocht er sterke samenhang zijn tussen de onafhankelijke variabelen, dan kan dat duiden op multicollineariteit, waardoor de coëfficiënten van die variabelen in de uiteindelijke modellen kan worden onderschat. Omdat leeftijd de enige continue variabele is en daarmee andere analysemethodes vereist, behandelen we die als laatst.

Passieve politieke participatie

```
CROSSTABS  
/TABLES=Passieve BY Actieve  
/FORMAT=AVALUE TABLES  
/STATISTICS=PHI  
/CELLS=COUNT  
/COUNT ROUND CELL.
```

```
CROSSTABS  
/TABLES=Passieve BY OPL werksituatie migratieachtergrond man  
/FORMAT=AVALUE TABLES
```

/STATISTICS=PHI
 /CELLS=COUNT
 /COUNT ROUND CELL.

Passieve en actieve politieke participatie

Wanneer we kijken naar de tabellen hieronder dan zien we dat de meeste respondenten wel doen aan passieve politieke participatie, maar niet aan actieve. De kleinste groep is die waar de respondenten wel doen aan actieve politieke participatie, maar niet aan passieve politieke participatie. De *Phi*-coëfficiënt is 0,091, met $p < 0,001$. Er is dus niet een hele sterke samenhang, maar deze is wel significant.

Passieve * Actieve Crosstabulation

Count		Actieve		Total
		,00	1,00	
Passieve	,00	248	20	268
	1,00	1893	443	2336
Total		2141	463	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,091	<,001
	Cramer's V	,091	<,001
N of Valid Cases		2604	

Passieve politieke participatie en opleidingsniveau

Zoals verwacht is er een significant verband tussen opleidingsniveau en passieve politieke participatie ($r=0,142$, $p < 0,001$). Te zien aan de kruistabel is dat van hoogopgeleiden (94,1%) in verhouding tot de andere laag (83,9%) en gemiddeld (86,2%) de meeste respondenten passief politiek participeren.

Passieve * OPL Crosstabulation

Count		OPL			Total
		,00	1,00	2,00	
Passieve	,00	59	134	75	268
	1,00	307	839	1190	2336
Total		366	973	1265	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,142	<,001
	Cramer's V	,142	<,001
N of Valid Cases		2604	

Passieve politieke participatie en werksituatie

Zoals verwacht is er ook een significant verband tussen werksituatie en passieve politieke participatie ($r = 0,088$, $p < 0,001$). De mate waarin de verschillende groepen meedoen aan passieve politieke participatie verschillen dus significant: 92,8% van de werknemers in de overheidssector, 87,4% van de werknemers in de private sector en 92,5% van de zelfstandigen.

Passieve * werksituatie Crosstabulation

Count		werksituatie			Total
		,00	1,00	2,00	
Passieve	,00	65	188	15	268
	1,00	849	1303	184	2336
Total		914	1491	199	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,088	<,001
	Cramer's V	,088	<,001
N of Valid Cases		2604	

Passieve politieke participatie en migratieachtergrond

Wat betreft de samenhang tussen deze twee variabelen zal het geen verassing zijn dat de correlatie significant en vrij sterk is: $r=0,132$, $p<0,001$. 91,3% van de mensen zonder migratieachtergrond doet mee aan passieve politieke participatie, 78,3% van de eerste generatie immigranten doet mee en 85,7% van de tweede generatie doen mee aan passieve politieke participatie.

Passieve * migratieachtergrond Crosstabulation

Count

		migratieachtergrond			Total
		,00	1,00	2,00	
Passieve	,00	187	43	38	268
	1,00	1968	139	229	2336
Total		2155	182	267	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,132	<,001
	Cramer's V	,132	<,001
N of Valid Cases		2604	

Passieve politieke participatie en geslacht

Het is opvallend dat het verschil in passieve politieke participatie tussen mannen en vrouwen miniem is ($Phi=0,001$, $p=0,956$). In tegenstelling tot wat verwacht was doen mannen en vrouwen dus ongeveer evenveel aan passieve politieke participatie.

Passieve * man Crosstabulation

Count

		man		Total
		,00	1,00	
Passieve	,00	144	124	268
	1,00	1251	1085	2336
Total		1395	1209	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,001	,956
	Cramer's V	,001	,956
N of Valid Cases		2604	

Actieve politieke participatie

CROSSTABS

/TABLES=Actieve BY OPL werksituatie migratieachtergrond man

/FORMAT=AVALUE TABLES

/STATISTICS=PHI

/CELLS=COUNT

/COUNT ROUND CELL.

Actieve politieke participatie en opleidingsniveau

Net als bij passieve politieke participatie is er bij actieve sprake van een significante samenhang met opleidingsniveau ($r = 0,124$, $p < 0,001$). Dat betekent dat er significante verschillen zijn in de verhoudingen binnen de groepen van opleidingsniveau waarop die meedoen aan actieve politieke participatie: laag 10,4%, gemiddeld 14,5% en hoog 22,5%.

Crosstab

Count		OPL			Total
		,00	1,00	2,00	
Actieve	,00	328	832	981	2141
	1,00	38	141	284	463
Total		366	973	1265	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,124	<,001
	Cramer's V	,124	<,001
N of Valid Cases		2604	

Actieve politieke participatie en werksituatie

We zien dat er significante verschillen zijn in actieve politieke participatie over de verschillende groepen van werksituatie ($r = 0,071$, $p = 0,001$): werknemers uit de overheidssector 20,1%, werknemers uit de private sector 15,6% en zelfstandigen 23,6%.

Crosstab

Count		werksituatie			Total
		,00	1,00	2,00	
Actieve	,00	730	1259	152	2141
	1,00	184	232	47	463
Total		914	1491	199	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,071	,001
	Cramer's V	,071	,001
N of Valid Cases		2604	

Actieve politieke participatie en migratieachtergrond

In tegenstelling tot wat verwacht is, zijn er geen significante verschillen in actieve politieke participatie op basis van de migratieachtergrond ($r = 0,039$, $p = 0,132$). Dat betekent dat migranten uit de eerste generatie, de tweede generatie en mensen zonder migratieachtergrond ongeveer in dezelfde verhouding meedoen aan actieve politieke participatie.

Crosstab

Count		migratieachtergrond			Total
		,00	1,00	2,00	
Actieve	,00	1785	148	208	2141
	1,00	370	34	59	463
Total		2155	182	267	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,039	,132
	Cramer's V	,039	,132
N of Valid Cases		2604	

Actieve politieke participatie en geslacht

In tegenstelling tot bij passieve politieke participatie zijn er wel significante verschillen tussen mannen en vrouwen in de verhouding dat zij meedoen aan actieve politieke participatie

($\Phi=0,073$, $p<0,001$). In verhouding doen meer mannelijke respondenten (20,7%) mee aan actieve politieke participatie dan vrouwen (15,2%).

Crosstab

Count

		man		Total
		,00	1,00	
Actieve	,00	1183	958	2141
	1,00	212	251	463
Total		1395	1209	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,073	<,001
	Cramer's V	,073	<,001
N of Valid Cases		2604	

Opleidingsniveau

CROSSTABS

```

/TABLES=OPL BY werksituatie migratieachtergrond man
/FORMAT=AVALUE TABLES
/STATISTICS=PHI
/CELLS=COUNT
/COUNT ROUND CELL.
    
```

Opleidingsniveau en werksituatie

Tussen opleidingsniveau en werksituatie is een significante en redelijk sterke samenhang: $r=0,138$, $p<0,001$. Wanneer we naar de kruistabel zien, dan zien we dat vooral de verdeling van hoogopgeleiden over de werksituaties sterk anders is dan die van laag en gemiddeld: hoog (43,4% overheidssector, 47,5% private sector, 9% zelfstandige), gemiddeld (28,8% overheidssector, 65,1% private sector en 6,2% zelfstandige) en laag (23,2% overheidssector, 70,2% private sector en 6,6% zelfstandige). Dit kan duiden op multicollineariteit, dus het is van belang om bij de modevaluatie in bijlage 3 daar goed naar te kijken of het de coëfficiënten te veel beïnvloed.

Crosstab

Count

		werksituatie			Total
		,00	1,00	2,00	
OPL	,00	85	257	24	366
	1,00	280	633	60	973
	2,00	549	601	115	1265
Total		914	1491	199	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,195	<,001
	Cramer's V	,138	<,001
N of Valid Cases		2604	

Opleidingsniveau en migratieachtergrond

Er is geen significante samenhang tussen opleidingsniveau en migratieachtergrond ($r = 0,020$, $p = 0,701$). Dat betekent dat er geen significante verschillen in opleidingsniveau over de migratieachtergronden is.

Crosstab

Count

		migratieachtergrond			Total
		,00	1,00	2,00	
OPL	,00	306	25	35	366
	1,00	810	60	103	973
	2,00	1039	97	129	1265
Total		2155	182	267	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,029	,701
	Cramer's V	,020	,701
N of Valid Cases		2604	

Opleidingsniveau en geslacht

Er is daarnaast ook geen significant verband tussen opleidingsniveau en geslacht ($r = 0,019$, $p = 0,626$), dus mannen en vrouwen hebben in deze steekproef ongeveer hetzelfde opleidingsniveau.

Crosstab

Count		man		Total
		,00	1,00	
OPL	,00	188	178	366
	1,00	521	452	973
	2,00	686	579	1265
Total		1395	1209	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,019	,626
	Cramer's V	,019	,626
N of Valid Cases		2604	

Werksituatie**CROSSTABS**

```

/TABLES=werksituatie BY migratieachtergrond man
/FORMAT=AVALUE TABLES
/STATISTICS=PHI
/CELLS=COUNT
/COUNT ROUND CELL.

```

Werksituatie en migratieachtergrond

Het verband tussen werksituatie en migratieachtergrond is niet significant ($r = 0,043$, $p = 0,047$).

Van de mensen zonder migratieachtergrond werkt 36,1% bij de overheid of een semi-overheid, 56,0% in de private sector en 13,6% als zelfstandige. Van de eerste generatie immigranten werken 28,6% in de overheidssector, 66,5% in de private sector en 4,9% als zelfstandige. Tot slot werken van de tweede generatie immigranten 31,2% in de overheidssector, 61,0% in de private sector en 7,1% als zelfstandige.

Crosstab

Count		migratieachtergrond			Total
		,00	1,00	2,00	
werksituatie	,00	777	52	85	914
	1,00	1207	121	163	1491
	2,00	171	9	19	199
Total		2155	182	267	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,061	,047
	Cramer's V	,043	,047
N of Valid Cases		2604	

Werksituatie en geslacht

Er zijn significante verschillen in werksituatie tussen mannen en vrouwen ($r=0,152$, $p<0,001$). Mannen werken vaker dan vrouwen in de overheidssector (41,6% ten opzichte van 27,6%), terwijl vrouwen vaker in de private sector werken (65,0% ten opzichte van 50,5%). Mannen werken daarnaast ook wat vaker als zelfstandige (7,9% tegenover 7,4%).

Crosstab

Count		man		Total
		,00	1,00	
werksituatie	,00	580	334	914
	1,00	705	786	1491
	2,00	110	89	199
Total		1395	1209	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,152	<,001
	Cramer's V	,152	<,001
N of Valid Cases		2604	

Migratieachtergrond

CROSSTABS

```

/TABLES=migratieachtergrond BY man
/FORMAT=AVALUE TABLES
/STATISTICS=PHI
/CELLS=COUNT
/COUNT ROUND CELL.
    
```

Migratieachtergrond en geslacht

Er zijn geen significante verschillen in geslacht over de groepen van migratieachtergrond ($r = 0,043$, $p = 0,093$).

migratieachtergrond * man Crosstabulation

Count		man		Total
		,00	1,00	
migratieachtergrond	,00	1165	990	2155
	1,00	103	79	182
	2,00	127	140	267
Total		1395	1209	2604

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,043	,093
	Cramer's V	,043	,093
N of Valid Cases		2604	

Leeftijd

CORRELATIONS

```

/VARIABLES=lft Passieve Actieve man
/PRINT=TWOTAIL NOSIG FULL
/MISSING=PAIRWISE.
    
```

ONEWAY lft BY opl

```

/MISSING ANALYSIS.
    
```

MEANS TABLES=lft BY Opl

```

/CELLS=MEAN COUNT STDDEV.
    
```

ONEWAY lft BY werksituatie

```

/MISSING ANALYSIS.
    
```

MEANS TABLES=lft BY werksituatie

```

/CELLS=MEAN COUNT STDDEV.
    
```

ONEWAY lft BY migratieachtergrond

/MISSING ANALYSIS.

MEANS TABLES=Ift BY migratieachtergrond
/CELLS=MEAN COUNT STDDEV.**Leeftijd met passieve en actieve politieke participatie en geslacht**

Leeftijd heeft een significante positieve relatie met passieve politieke participatie ($r=0,082$, $p<0,001$). Dat betekent dat naar mate mensen ouder worden ze eerder mee gaan doen aan passieve politieke participatie. Maar het verband is niet heel sterk: mensen die wel passief participeren zijn maar 0,082 jaar ouder. Het is opvallen dat datzelfde effect niet optreedt bij actieve politieke participatie: $r=0,017$, $p=0,374$. Tot slot is er een significante relatie tussen leeftijd en geslacht ($r=0,094$, $p<0,001$). Gezien de correlatiecoëfficiënt positief is, betekent dat de mannen in de steekproef gemiddeld ouder zijn dan vrouwen.

Correlations

		Ift	Passieve	Actieve	man
Ift	Pearson Correlation	1	,082**	,017	,094**
	Sig. (2-tailed)		<,001	,374	<,001
	N	2604	2604	2604	2604
Passieve	Pearson Correlation	,082**	1	,091**	,001
	Sig. (2-tailed)	<,001		<,001	,956
	N	2604	2604	2604	2604
Actieve	Pearson Correlation	,017	,091**	1	,073**
	Sig. (2-tailed)	,374	<,001		<,001
	N	2604	2604	2604	2604
man	Pearson Correlation	,094**	,001	,073**	1
	Sig. (2-tailed)	<,001	,956	<,001	
	N	2604	2604	2604	2604

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

Leeftijd en opleidingsniveau

Er is een significant verschil gemiddelde in leeftijd tussen de verschillende opleidingsniveaus ($F=41,6$, $df=2$, $p<0,001$). We zien dat laagopgeleiden in de steekproef gemiddeld 5,7 jaar ouder zijn dan het gemiddelde, terwijl hoogopgeleiden 1,7 jaar jonger zijn dan het gemiddelde.

ANOVA

lft	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	15625,351	2	7812,676	41,623	<,001
Within Groups	488208,327	2601	187,700		
Total	503833,679	2603			

Report

lft	Mean	N	Std. Deviation
OPL			
,00	3,8486	366	13,91017
1,00	-1,9874	973	13,43287
2,00	-3,5696	1265	13,84229
Total	-1,9357	2604	13,91254

Leeftijd en werksituatie

Daarnaast zijn er ook significante verschillen in de gemiddelde leeftijd over de verschillende werksituaties ($F=32,1$, $df=2$, $p<0,001$). Zelfstandigen zijn 5,4 jaar ouder dan het gemiddelde, terwijl werknemers uit de private sector 1,7 jaar jonger zijn dan het gemiddelde. Mensen uit de overheidssector zijn gemiddeld 1,5 jaar jonger dan het gemiddelde.

ANOVA

lft	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	12148,200	2	6074,100	32,132	<,001
Within Groups	491685,478	2601	189,037		
Total	503833,679	2603			

Report

lft	Mean	N	Std. Deviation
werksituatie			
,00	-,4208	914	13,65279
1,00	-3,5942	1491	13,85177
2,00	3,5322	199	13,41167
Total	-1,9357	2604	13,91254

Leeftijd en migratieachtergrond

Tot slot zijn er ook significante verschillen in leeftijd over de verschillende migratieachtergronden ($F=24,1$, $df=2$, $p<0,001$). Terwijl de gemiddelde leeftijden van respondenten zonder migratieachtergrond en van de eerste generatie immigranten redelijk dicht bij het gemiddelde van de steekproef zitten, zijn de tweede generatie immigranten gemiddeld veel jonger: 5,6 jaar.

ANOVA

lft	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	9172,209	2	4586,104	24,114	<,001
Within Groups	494661,470	2601	190,181		
Total	503833,679	2603			

Report

lft	Mean	N	Std. Deviation
migratieachtergrond			
,00	-1,3065	2155	13,97318
1,00	-1,2407	182	11,77761
2,00	-7,4880	267	13,57057
Total	-1,9357	2604	13,91254

Multivariate statistieken

Om onze hypothesen te toetsen schatten we vier verschillende binaire logistische regressie modellen. Voor modellen A en B zijn de onafhankelijke variabelen: laag opleidingsniveau, hoog opleidingsniveau, zelfstandige, werknemer private sector, eerste generatie, tweede generatie, man, leeftijd en interactievariabelen tussen zelfstandige en de dummy's van opleidingsniveau. Deze zijn niet handmatig gecodeerd, maar in de modellen als product van elkaar geplott.

In modellen C en D is er dus extra geselecteerd op hoogopgeleiden, waardoor er geen variabelen van opleidingsniveau als onafhankelijke variabele zijn opgenomen. Daardoor zijn ook de interactievariabelen uit deze modellen.

Model	Hypothesen	Extra selectie	Afhankelijke variabele
A	1 & 5		Passieve politieke participatie
B	2 & 5		Actieve politieke participatie

C	3	Hoog opgeleid	Passieve politieke participatie
D	4	Hoog opgeleid	Actieve politieke participatie

Hieronder bespreken we de geschatte coëfficiënten van de regressiemodellen en wat die inhoudelijk betekenen. Daarbij benoemen we de helling, de *p*-waarde en de odds-ratio. De helling laat zien wat de verandering in odds-ratio is wanneer we één extra eenheid bij de onafhankelijke variabele optellen. De *p*-waarde laat zien hoe groot de kans is dat die helling significant verschilt van 0. Tot slot bespreken we de odds-ratio's; deze kunnen we interpreteren als de waarde waarmee de odds worden vermenigvuldigd, wanneer de onafhankelijke variabele met één eenheid stijgt. Voor een constante staat de coëfficiënt voor het gemiddelde. We bespreken voor de constante niet de odds-ratio, omdat deze niet met één eenheid kan stijgen.

Met deze waarden doen we in de resultatenparagraaf uitspraken over de hypothesen. In bijlage 3 evalueren we de kwaliteit van de modellen.

Model A: Passieve politieke participatie

```
LOGISTIC REGRESSION VARIABLES Passieve
/METHOD=ENTER lft man Gen1 Gen2 wnpriv
/METHOD=ENTER Zelfstandige OPLlaag OPLhoog
/METHOD=ENTER OPLlaag*Zelfstandige OPLhoog*Zelfstandige
/CLASSPLOT
/PRINT=GOODFIT CI(98)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

Model A schat passieve politieke participatie. In Blok 0 is daarvoor alleen de constante gebruikt. Die schat daarmee een coëfficiënt van: $b=2,165$ en $p<0,001$. In blok 1 zijn daar de controlevariabelen aan toegevoegd. Elk daarvan is significant, behalve man: $b=0,034$, $p=0,802$ en odds-ratio=1,034. Dat betekent dat er geen significant verschil is in passieve politieke participatie tussen mannen en vrouwen, gecontroleerd voor de andere variabelen. De andere controlevariabelen hebben wel significante coëfficiënten en ook in de richting die werd verwacht: leeftijd ($b=0,017$, $p<0,001$, odds-ratio=1,017), eerste generatie $b=-1,153$, $p<0,001$, odds-ratio=0,316), tweede generatie ($b=-0,444$, $p=0,022$ en odds-ratio=0,641) en werknemer uit de private sector ($b=-0,527$, $p<0,001$, odds-ratio=0,590). De constante is in dit model gestegen naar $b=2,712$ met $p<0,001$.

In blok 2 zijn ook de onafhankelijke variabelen toegevoegd. Geen van de controle variabelen is daardoor verandert in significantie, maar zoals je kon verwachten is wel de coëfficiënt van

werknemers uit de private sector sterk verandert: $b=-0,353$, $p=0,025$ en odds-ratio=0,702). Van de onafhankelijke variabelen is alleen hoog opleidingsniveau significant: $b=0,964$, $p<0,001$, odds-ratio=2,622). Laag opleidingsniveau ($b=-0,300$, $p=0,089$ en odds-ratio=0,741) en zelfstandige ($b=-0,142$, $p=0,641$ en odds-ratio=0,868) zijn niet significant.

In blok 3 zijn tenslotte de interacties toegevoegd. Deze zijn beide niet significant en hebben bovendien zeer kleine coëfficiënten: laagopgeleid*zelfstandige ($b=-0,077$, $p=0,920$ en odds-ratio=0,926) en hoog opgeleid*zelfstandige ($b=0,080$, $p=0,900$ en odds-ratio=1,083). Het is dan ook niet verassend dat de coëfficiënten van de andere variabelen zeer weinig zijn veranderd.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	2,165	,064	1127,109	1	<,001	8,716

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	98% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	lft	,017	,005	11,626	1	<,001	1,017	1,005	1,028
	man	,034	,134	,063	1	,802	1,034	,758	1,411
	Gen1	-1,153	,193	35,800	1	<,001	,316	,202	,494
	Gen2	-,444	,194	5,225	1	,022	,641	,408	1,008
	wnpriv	-,527	,144	13,423	1	<,001	,590	,423	,825
	Constant	2,712	,133	418,648	1	<,001	15,052		

a. Variable(s) entered on step 1: lft, man, Gen1, Gen2, wnpriv.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	98% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	lft	,023	,005	21,515	1	<,001	1,023	1,012	1,035
	man	,023	,135	,028	1	,867	1,023	,748	1,399
	Gen1	-1,256	,198	40,265	1	<,001	,285	,180	,451
	Gen2	-,434	,198	4,837	1	,028	,648	,409	1,025
	wnpriv	-,353	,157	5,044	1	,025	,702	,487	1,013
	Zelfstandige	-,142	,304	,217	1	,641	,868	,428	1,761
	OPLlaag	-,300	,176	2,887	1	,089	,741	,492	1,117
	OPLhoog	,964	,155	38,470	1	<,001	2,622	1,827	3,764
	Constant	2,324	,163	202,082	1	<,001	10,216		

a. Variable(s) entered on step 1: Zelfstandige, OPLlaag, OPLhoog.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	98% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	lft	,023	,005	21,484	1	<,001	1,023	1,011	1,035
	man	,023	,135	,029	1	,864	1,023	,748	1,400
	Gen1	-1,256	,198	40,194	1	<,001	,285	,180	,452
	Gen2	-,435	,198	4,844	1	,028	,647	,409	1,025
	wnpriv	-,355	,157	5,070	1	,024	,702	,486	1,012
	Zelfstandige	-,157	,437	,129	1	,720	,855	,309	2,363
	OPLlaag	-,295	,181	2,658	1	,103	,744	,488	1,134
	OPLhoog	,958	,161	35,602	1	<,001	2,608	1,795	3,789
	OPLlaag by Zelfstandige	-,077	,765	,010	1	,920	,926	,156	5,493
	OPLhoog by Zelfstandige	,080	,636	,016	1	,900	1,083	,247	4,757
	Constant	2,325	,164	200,345	1	<,001	10,231		

a. Variable(s) entered on step 1: OPLlaag * Zelfstandige , OPLhoog * Zelfstandige .

Model B: actieve politieke participatie

```
LOGISTIC REGRESSION VARIABLES Actieve
/METHOD=ENTER lft man Gen1 Gen2 wnpriv
/METHOD=ENTER Zelfstandige OPLlaag OPLhoog
/METHOD=ENTER OPLlaag*Zelfstandige OPLhoog*Zelfstandige
/CLASSPLOT
/PRINT=GOODFIT CI(98)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

In het model zonder variabelen is de constante -1,531 met $p < 0,001$. Dat deze constante veel lager is dan die van passieve politieke participatie is logisch doordat de gemiddelde kans op actieve politieke lager is dan die van passieve.

In blok 1 zijn de controlevariabelen toegevoegd en opvallend is dat er andere variabelen significant zijn, dan in het model van passieve participatie. Zo is man deze keer wel significant ($b=0,435$, $p < 0,001$ en odds-ratio=1,546), wat betekent dat de kans dat een man gaat meedoen aan actieve politieke participatie significant groter is dan dat een vrouw dat gaat doen. Daarnaast is de coëfficiënt van werknemer uit de private sector significant negatief: $b=-0,425$, $p < 0,001$ en odds-ratio=0,654). De andere coëfficiënten zijn niet significant: leeftijd ($b=0,001$, $p=0,850$, odds-ratio=0,850), eerste generatie ($b=0,162$, $p=0,417$, odds-ratio=1,176), tweede generatie ($b=0,310$, $p=0,055$ en odds-ratio=1,386). Opvallend is hieraan dat hoewel ze niet significant zijn, eerste en tweede generatie migranten een positieve coëfficiënt hebben, wat betekent dat de kans groter is dat zij meedoen aan actieve politieke participatie dan respondenten zonder migratieachtergrond. Dit terwijl in het model van passieve politieke participatie ze significante negatieve coëfficiënten hebben.

In blok 2 zijn zelfstandigen en de opleidingsniveaus toegevoegd. Daarvan is alleen hoog opleidingsniveau significant: $b=0,964$, $p < 0,001$ en odds-ratio=2,622. Hoog opleidingsniveau heeft dus een sterk effect op actieve politieke participatie. Maar laag opleidingsniveau heeft geen significante coëfficiënt: $b=-0,300$, $p=0,089$, odds-ratio=0,741. Daarnaast heeft ook zelfstandige geen significante coëfficiënt: $b=-0,142$, $p=0,641$ en odds-ratio=0,868.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-1,531	,051	892,643	1	<,001	,216

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	98% C.I.for EXP(B)	
							Lower	Upper
Step 1 ^a lft	,001	,004	,036	1	,850	1,001	,992	1,010
man	,435	,106	17,013	1	<,001	1,546	1,209	1,976
Gen1	,162	,200	,659	1	,417	1,176	,739	1,874
Gen2	,310	,161	3,686	1	,055	1,363	,936	1,985
wnpriv	-,425	,106	16,051	1	<,001	,654	,511	,837
Constant	-1,560	,090	299,551	1	<,001	,210		

a. Variable(s) entered on step 1: lft, man, Gen1, Gen2, wnpriv.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	98% C.I.for EXP(B)	
							Lower	Upper
Step 1 ^a lft	,004	,004	1,200	1	,273	1,004	,995	1,014
man	,420	,107	15,462	1	<,001	1,521	1,187	1,950
Gen1	,127	,202	,399	1	,528	1,136	,711	1,815
Gen2	,335	,163	4,240	1	,039	1,397	,958	2,040
wnpriv	-,261	,115	5,148	1	,023	,770	,589	1,007
Zelfstandige	,179	,189	,900	1	,343	1,197	,771	1,858
OPLlaag	-,404	,196	4,235	1	,040	,668	,423	1,054
OPLhoog	,500	,116	18,679	1	<,001	1,649	1,260	2,158
Constant	-1,873	,128	214,728	1	<,001	,154		

a. Variable(s) entered on step 1: Zelfstandige, OPLlaag, OPLhoog.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	98% C.I.for EXP(B)	
							Lower	Upper
Step 1 ^a lft	,004	,004	1,123	1	,289	1,004	,995	1,013
man	,426	,107	15,898	1	<,001	1,532	1,194	1,964
Gen1	,134	,202	,439	1	,508	1,143	,715	1,828
Gen2	,338	,163	4,312	1	,038	1,402	,960	2,047
wnpriv	-,270	,115	5,493	1	,019	,763	,583	,998
Zelfstandige	-,205	,383	,287	1	,592	,815	,334	1,986
OPLlaag	-,421	,204	4,261	1	,039	,657	,409	1,055
OPLhoog	,454	,121	14,130	1	<,001	1,575	1,189	2,086
OPLlaag by Zelfstandige	,264	,746	,125	1	,723	1,302	,230	7,384
OPLhoog by Zelfstandige	,540	,434	1,545	1	,214	1,715	,625	4,710
Constant	-1,844	,129	203,649	1	<,001	,158		

a. Variable(s) entered on step 1: OPLlaag * Zelfstandige, OPLhoog * Zelfstandige.

Model C: Passieve politieke participatie geselecteerd op hoog opgeleid

LOGISTIC REGRESSION VARIABLES Passieve
 /METHOD=ENTER lft man Gen1 Gen2 wnpriv
 /METHOD=ENTER Zelfstandige
 /PRINT=GOODFIT CI(95)
 /CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).

In dit model is passieve politieke participatie de afhankelijke variabele. In tegenstelling tot model A is er echter geselecteerd op mensen die hoogopgeleid zijn. De steekproef omvang van dit model is daarom: 1265.

In het nul model is de constante 2,764 met $p < 0,001$. In stap 1 zijn de controlevariabelen toegevoegd. Daarvan zijn eerste generatie ($b = -1,734$, $p < 0,001$, odds-ratio = 0,177) en tweede generatie ($b = -0,884$, $p = 0,007$, odds-ratio = 0,217) significant. De negatieve waarde betekent respondenten met een migratieachtergrond significant minder aan passieve politieke participatie doen. Het effect van leeftijd is ook significant, maar zeer klein ($b = 0,029$, $p = 0,004$). Opvallend is dat het verschil tussen man en vrouw miniem is: $b = 0,002$, $p = 0,994$ en odds-ratio = 1,002. Daarnaast is het opvallend dat voor hoogopgeleiden de coëfficiënt van werknemers uit de private sector niet significant is: $b = -0,263$, $p = 0,308$ en odds-ratio = 0,769.

In stap 2 komt ten slotte zelfstandige daar nog bij als variabele, deze is niet significant: $b = -0,101$, $p = 0,843$ en odds-ratio = 0,904. Wanneer we dit model vergelijken met het model waar alle opleidingsniveaus zijn inbegrepen, dan valt op dat de coëfficiënten van eerste en tweede generatie veel lager zijn voor hoger opgeleiden.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	2,764	,119	539,092	1	<,001	15,867

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	lft	,029	,010	8,461	1	,004	1,029	1,010	1,050
	man	,002	,255	,000	1	,994	1,002	,608	1,651
	Gen1	-1,734	,313	30,684	1	<,001	,177	,096	,326
	Gen2	-,884	,328	7,282	1	,007	,413	,217	,785
	wnpriv	-,263	,258	1,041	1	,308	,769	,464	1,274
	Constant	3,446	,238	210,408	1	<,001	31,385		

a. Variable(s) entered on step 1: lft, man, Gen1, Gen2, wnpriv.

		Variables in the Equation					95% C.I. for EXP(B)		
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1 ^a	lft	,029	,010	8,495	1	,004	1,030	1,010	1,050
	man	,002	,255	,000	1	,993	1,002	,608	1,651
	Gen1	-1,736	,313	30,715	1	<,001	,176	,095	,326
	Gen2	-,882	,328	7,233	1	,007	,414	,218	,787
	wnpriv	-,279	,271	1,060	1	,303	,756	,445	1,287
	Zelfstandige	-,101	,510	,039	1	,843	,904	,332	2,457
	Constant	3,463	,254	185,984	1	<,001	31,928		

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

Model D: Actieve politieke participatie geselecteerd op hoog opleidingsniveau

LOGISTIC REGRESSION VARIABLES Actieve
 /METHOD=ENTER lft man Gen1 Gen2 wnpriv
 /METHOD=ENTER Zelfstandige
 /PRINT=GOODFIT CI(95)
 /CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).

In dit model is ook geselecteerd op hoogopgeleiden, maar is actieve politieke participatie de afhankelijke variabele. De constante van het nul-model is -1,240 met $p < 0,001$.

Met stap 1 zijn de controle variabelen weer toegevoegd en daarbij zijn in tegenstelling tot passieve politieke participatie juist man ($b=0,434$, $p=0,002$, odds-ratio=1,544) en werknemers uit de private sector ($b=-0,354$, $p=0,015$, odds-ratio=0,702) significant. De controlevariabelen die bij passieve participatie significant zijn, zijn dat nu juist niet: leeftijd ($b=0,009$, $p=0,078$, odds-ratio=1,009), eerste generatie ($b=0,146$, $p=0,565$, odds-ratio=1,157) en tweede generatie ($b=0,121$, $p=0,592$, odds-ratio=1,129).

In stap 2 is zelfstandige toegevoegd. Net als in het model van passieve politieke participatie is de coëfficiënt van zelfstandige niet significant: $b=0,286$, $p=0,210$ en odds-ratio=1,332. Wanneer we dit laatste model vergelijken met het uiteindelijke model van actieve politieke participatie zonder extra selectie, dan valt op dat daar tweede generatie een significante coëfficiënt heeft en hier niet.

		Variables in the Equation					Exp(B)
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-1,240	,067	338,422	1	<,001	,290

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step 1 ^a lft	,009	,005	3,113	1	,078	1,009	,999	1,019
man	,434	,142	9,379	1	,002	1,544	1,169	2,038
Gen1	,146	,254	,332	1	,565	1,157	,704	1,903
Gen2	,121	,226	,286	1	,592	1,129	,725	1,757
wnpriv	-,354	,145	5,916	1	,015	,702	,528	,934
Constant	-1,288	,111	134,392	1	<,001	,276		

a. Variable(s) entered on step 1: lft, man, Gen1, Gen2, wnpriv.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step 1 ^a lft	,009	,005	2,705	1	,100	1,009	,998	1,019
man	,434	,142	9,362	1	,002	1,543	1,169	2,038
Gen1	,153	,254	,364	1	,546	1,166	,709	1,918
Gen2	,117	,226	,269	1	,604	1,124	,722	1,751
wnpriv	-,304	,151	4,023	1	,045	,738	,549	,993
Zelfstandige	,286	,229	1,568	1	,210	1,332	,851	2,085
Constant	-1,341	,120	125,067	1	<,001	,262		

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

Bijlage 3 Modevaluatie

In deze bijlage beoordelen we de geschatte regressiemodellen op kwaliteit. Dat doen we door te kijken naar de assumpties van binaire logistische regressie, eventuele *outliers* en invloedrijke punten, multicollineariteit en tot slot de *fit* van de modellen. Aan de hand hiervan kunnen we beter beoordelen wat de uitkomsten van het onderzoek betekenen.

Assumpties

Voor binaire logistische regressie wordt er maar één assumptie gedaan. Deze assumptie is dat er sprake is van onafhankelijke steekproeftrekking. Dit betekent dat de respondenten geen invloed op elkaar uitoefenen. Bij deze data wordt er niet helemaal aan deze assumptie voldaan; dat komt omdat er meerdere personen uit hetzelfde huishouden in de steekproef kunnen zitten. Het kan daardoor voorkomen dat die respondenten de scores van elkaar en dus de resultaten van de analyses beïnvloeden. Het gevolg van deze schending is dat de daadwerkelijke steekproefomvang is overschat. Doordat deze gebruikt wordt om de *p*-waardes en de standaardfouten te schatten, worden ook deze overschat. Om hiervoor te controleren gebruiken we een significantieniveau van 0,01 in plaats van 0,05.

Outliers en invloedrijke punten

Outliers en invloedrijke punten zijn respondenten met scores die sterk afwijken van de rest en daardoor een relatief grote invloed hebben op de uiteindelijke modellen. Doordat de steekproef een redelijk grote omvang heeft en er maar één continue variabele is, is de kans op *outliers* en invloedrijke punten kleiner. Om er zeker van te zijn dat er geen sprake van is, controleren we toch voor deze punten. Dat doen we door de vier uiteindelijke regressiemodellen opnieuw te schatten, waarbij de DFBETA en de gestandaardiseerde residuen van de respondenten zijn opgeslagen.

De DFBETA is de mate waarin een respondent van invloed is op de coëfficiënt van een variabele. Als kritieke waarde wordt daarvoor aangehouden: $\frac{2}{\sqrt{n}}$. Voor modellen A en B ($n=2604$) is dat 0,039, terwijl voor C en D ($n=1265$) deze kritieke waarde 0,056 bedraagt. Daarnaast kijken we naar de gestandaardiseerde residuen. Een te groot gestandaardiseerd residu betekent dat de geschatte kans van de afhankelijke variabele te ver afwijkt van de daadwerkelijke waarde. Daarbij houden we aan dat deze tussen -3 en 3 moet zitten. Hoger dan 3, dan wordt de kans overschat; lager dan 3, dan wordt de kans onderschat.

```
LOGISTIC REGRESSION VARIABLES Passieve
/METHOD=ENTER Zelfstandige OPLlaag OPLhoog wnpriv Gen1 Gen2 lft man
OPLlaag*Zelfstandige
```

```
OPLhoog*Zelfstandige
/SAVE=DFBETA ZRESID
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

```
LOGISTIC REGRESSION VARIABLES Actieve
/METHOD=ENTER Zelfstandige OPLlaag OPLhoog wnpriv Gen1 Gen2 lft man
OPLlaag*Zelfstandige
OPLhoog*Zelfstandige
/SAVE=DFBETA ZRESID
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

```
COMPUTE filter_$=(OPLhoog = 1 & Coo_1>0).
VARIABLE LABELS filter_$ 'OPLhoog = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

```
LOGISTIC REGRESSION VARIABLES Passieve
/METHOD=ENTER Zelfstandige wnpriv Gen1 Gen2 lft man
/SAVE=DFBETA ZRESID
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

```
LOGISTIC REGRESSION VARIABLES ACTIEVE
/METHOD=ENTER Zelfstandige wnpriv Gen1 Gen2 lft man
/SAVE=DFBETA ZRESID
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

```
FREQUENCIES VARIABLES=ZRE_1 DFB0_1 DFB1_1 DFB2_1 DFB3_1 DFB4_1 DFB5_1
DFB6_1 DFB7_1 DFB8_1 DFB9_1
DFB10_1
/STATISTICS=MINIMUM MAXIMUM MEAN
/FORMAT=NOTABLE
/ORDER=ANALYSIS.
```

```
FREQUENCIES VARIABLES=ZRE_2 DFB0_2 DFB1_2 DFB2_2 DFB3_2 DFB4_2 DFB5_2
DFB6_2 DFB7_2 DFB8_2 DFB9_2
DFB10_2
/STATISTICS=MINIMUM MAXIMUM MEAN
/FORMAT=NOTABLE
/ORDER=ANALYSIS.
```

```
FREQUENCIES VARIABLES=ZRE_3 DFB0_3 DFB1_3 DFB2_3 DFB3_3 DFB4_3 DFB5_3
DFB6_3
/STATISTICS=MINIMUM MAXIMUM MEAN
/FORMAT=NOTABLE
/ORDER=ANALYSIS.
```

```
FREQUENCIES VARIABLES=ZRE_6 DFB0_6 DFB1_6 DFB2_6 DFB3_6 DFB4_6 DFB5_6
DFB6_6
/STATISTICS=MINIMUM MAXIMUM MEAN
/FORMAT=NOTABLE
/ORDER=ANALYSIS.
```

Model A

J.G. DE VRIES – POLITICAL PARTICIPATION IN THE LIGHT OF THE GOALFRAMING THEORY

Statistics

	Normalized residual	DFBETA for constant	DFBETA for Zelfstandige	DFBETA for OPLJaag	DFBETA for OPLhoog	DFBETA for wnpriv	DFBETA for Gen1	DFBETA for Gen2	DFBETA for lft	DFBETA for man	DFBETA for OPLJaag by Zelfstandige	DFBETA for OPLhoog by Zelfstandige
N	Valid Missing	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0
Mean		,0010322	,0000004	,0000008	,0000002	-,0000003	,0000000	-,0000001	,0000009	,0000000	-,0000008	-,0000024
Minimum		-6,28464	-,02733	-,15768	-,02219	-,01796	-,01121	-,03148	-,03169	-,00062	-,01193	-,35661
Maximum		,88775	,01274	,04897	,01230	,01144	,01906	,01516	,01130	,00071	,00969	,15969

Model B

Statistics

	Normalized residual	DFBETA for constant	DFBETA for Zelfstandige	DFBETA for OPLJaag	DFBETA for OPLhoog	DFBETA for wnpriv	DFBETA for Gen1	DFBETA for Gen2	DFBETA for lft	DFBETA for man	DFBETA for OPLJaag by Zelfstandige	DFBETA for OPLhoog by Zelfstandige
N	Valid Missing	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0	2604 0
Mean		-,0000899	,0000001	-,0000003	,0000000	,0000000	,0000001	-,0000002	-,0000001	,0000000	-,0000002	-,0000013
Minimum		-,89775	-,00594	-,02954	-,00963	-,01000	-,00976	-,01563	-,01122	-,00048	-,00739	-,11976
Maximum		3,55206	,01477	,12004	,03276	,00654	,00862	,03305	,02108	,00047	,00665	,35784

Model C

Statistics

	Normalized residual	DFBETA for constant	DFBETA for Zelfstandige	DFBETA for wnpriv	DFBETA for Gen1	DFBETA for Gen2	DFBETA for lft	DFBETA for man
N	Valid Missing	1265 0	1265 0	1265 0	1265 0	1265 0	1265 0	1265 0
Mean		-,0018036	,0000003	,0000024	,0000020	-,0000021	,0000015	,0000001
Minimum		-7,25030	-,07793	-,21160	-,04183	-,06911	-,08126	-,00253
Maximum		,69684	,02701	,05108	,05325	,03891	,03148	,00172

Model D

Statistics

	Normalized residual	DFBETA for constant	DFBETA for Zelfstandige	DFBETA for wnpriv	DFBETA for Gen1	DFBETA for Gen2	DFBETA for lft	DFBETA for man
N	Valid Missing	1265 0	1265 0	1265 0	1265 0	1265 0	1265 0	1265 0
Mean		,0006030	,0000001	,0000002	,0000002	-,0000003	,0000003	,0000000
Minimum		-,85152	-,00659	-,01801	-,01304	-,02640	-,02001	-,00061
Maximum		2,57452	,01198	,03560	,01622	,04808	,03845	,00082

Wanneer we deze modellen controleren op deze kritieke waarden, dan zien we dat alleen model D geen *outliers* en invloedrijke punten heeft. Model A heeft een te lage en te hoge DFBETA bij zelfstandige, respectievelijk -0,158 en 0,049. Gezien de interacties samenhangen met zelfstandige is het niet verassend dat hun DFBETA's ook te veel afwijken. Bovendien heeft model A een te lage waarde bij gestandaardiseerd residu: -6,3. Een negatief gestandaardiseerd residu, betekent dat de geschatte waarde van de respondenten hoger is dan de werkelijke waarde, dus dat de kans van de respondent op passieve politieke participatie is overschat. Voor model B geldt hetzelfde: een DFBETA van 0,120 bij zelfstandige en een gestandaardiseerd residu van maximaal 3,6. Hier is de kans op actieve politieke participatie dus onderschat. Model C heeft ook problemen met de DFBETA van zelfstandige: -0,212, maar daarnaast ook met eerste generatie (-0,069) en tweede generatie (-0,081). Opnieuw heeft ook model C een te laag minimum van gestandaardiseerde residuen; -7,3.

Er zijn dus wel degelijk *outliers* en invloedrijke punten. Om er zeker van te zijn dat deze niet een te grote invloed hebben schatten we de modellen opnieuw, waarbij deze respondenten eruit zijn gehaald. Maar eerst gaan we kijken wat voor respondenten dit zijn.

Model A

Gestandaardiseerd residuen

Model A heeft gestandaardiseerde residuen die lager zijn dan -3, wat betekent dat er respondenten zijn waarvan we hun geschatte kans op passieve politieke participatie overschat hebben. Er zijn 79 respondenten waar dat het geval is: zij hebben een 0-score bij passieve politieke participatie en hebben dus niet meegedaan aan passieve politieke participatie. Dit terwijl ze wel hoog scoren op onafhankelijke variabelen die veel invloed uitoefenen op de kans van participatie: veel hoogopgeleiden, veel mensen die werken in de overheidssector en geen migratieachtergrond.

```
USE ALL.
COMPUTE filter_$=(ZRE_1 < -3).
VARIABLE LABELS filter_$ 'ZRE_1 < -3 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

```
FREQUENCIES VARIABLES=Passieve OPL lft man migratieachtergrond werksituatie
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
/ORDER=ANALYSIS.
```

Passieve					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	79	100,0	100,0	100,0

OPL					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	3	3,8	3,8	3,8
	1,00	24	30,4	30,4	34,2
	2,00	52	65,8	65,8	100,0
	Total	79	100,0	100,0	

werksituatie

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	39	49,4	49,4	49,4
	1,00	31	39,2	39,2	88,6
	2,00	9	11,4	11,4	100,0
Total		79	100,0	100,0	

migratieachtergrond

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	70	88,6	88,6	88,6
	1,00	2	2,5	2,5	91,1
	2,00	7	8,9	8,9	100,0
Total		79	100,0	100,0	

man

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	47	59,5	59,5	59,5
	1,00	32	40,5	40,5	100,0
	Total		79	100,0	100,0

Statistics

		Passieve	OPL	lft	man	migratieachtergrond	werksituatie
N	Valid	79	79	79	79	79	79
	Missing	0	0	0	0	0	0
Mean		,0000	1,6203	-,3177	,4051	,2025	,6203
Std. Deviation		,00000	,56168	13,90592	,49404	,58572	,68508
Minimum		,00	,00	-26,90	,00	,00	,00
Maximum		,00	2,00	29,10	1,00	2,00	2,00

DFBETA

Een te lage DFBETA betekent dat deze respondenten in verhouding een grote invloed hebben op de coëfficiënt en dus de coëfficiënt omlaag trekken. Kijken we naar de respondenten dan zien we dat het zeven zelfstandigen betreft die niet hebben meegedaan aan passieve politieke participatie. Ze zijn daarnaast relatief oud, gemiddeld opgeleid, zonder migratieachtergrond en ouder dan gemiddeld. Op basis van die kenmerken zouden we dus verwachten dat ze wel zouden participeren.

Kijken we daarnaast naar de respondenten die juist een te hoge DFBETA hebben, dan zijn dat allen zelfstandigen die wel hebben meegedaan aan passieve politieke participatie. Daarnaast hebben ze een gemiddeld opleidingsniveau, een migratieachtergrond en er is één met een zeer

lage leeftijd. Op basis van deze kenmerken zou je dus kunnen verwachten dat deze respondenten juist niet zouden participeren.

```
USE ALL.
COMPUTE filter_$=(DFB1_1<-0.044).
VARIABLE LABELS filter_$ 'ZRE_1 > 3 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

```
FREQUENCIES VARIABLES=Passieve OPL lft man migratieachtergrond werksituatie
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
/ORDER=ANALYSIS.
```

Passieve

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ,00	7	100,0	100,0	100,0

OPL

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1,00	7	100,0	100,0	100,0

werksituatie

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 2,00	7	100,0	100,0	100,0

migratieachtergrond

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ,00	5	71,4	71,4	71,4
1,00	1	14,3	14,3	85,7
2,00	1	14,3	14,3	100,0
Total	7	100,0	100,0	

man

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ,00	2	28,6	28,6	28,6
1,00	5	71,4	71,4	100,0
Total	7	100,0	100,0	

lft					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	-18,90	1	14,3	14,3	14,3
	-3,90	1	14,3	14,3	28,6
	2,10	1	14,3	14,3	42,9
	4,10	1	14,3	14,3	57,1
	5,10	1	14,3	14,3	71,4
	19,10	1	14,3	14,3	85,7
	20,10	1	14,3	14,3	100,0
Total		7	100,0	100,0	

USE ALL.

COMPUTE filter_\$=(DFB1_1>0.044).

VARIABLE LABELS filter_\$ 'ZRE_1 > 3 (FILTER)'.
 VALUE LABELS filter_\$ 0 'Not Selected' 1 'Selected'.
 FORMATS filter_\$ (f1.0).
 FILTER BY filter_\$.
 EXECUTE.

FORMATS filter_\$ (f1.0).

FILTER BY filter_\$.

EXECUTE.

FREQUENCIES VARIABLES=Passieve OPL lft man migratieachtergrond werksituatie
 /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
 /ORDER=ANALYSIS.

Passieve					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	4	100,0	100,0	100,0

OPL					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	4	100,0	100,0	100,0

lft					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	-27,90	1	25,0	25,0	25,0
	1,10	1	25,0	25,0	50,0
	4,10	2	50,0	50,0	100,0
Total		4	100,0	100,0	

man

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	3	75,0	75,0	75,0
	1,00	1	25,0	25,0	100,0
Total		4	100,0	100,0	

migratieachtergrond

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	3	75,0	75,0	75,0
	2,00	1	25,0	25,0	100,0
Total		4	100,0	100,0	

werksituatie

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2,00	4	100,0	100,0	100,0

Model B

Gestandaardiseerde residuen

Opvallend is dat bij model B dezelfde respondenten de invloedrijke punten lijken te zijn. Het zijn er opnieuw 79 respondenten. Maar deze keer zijn de gestandaardiseerde residuen dus te hoog, wat betekent dat de geschatte kans van de respondenten hoger is dan de daadwerkelijke waarde. De respondenten zijn over het algemeen gemiddeld of hoogopgeleid en ze werken vaak in de overheidssector of als werknemer in de private sector.

```
VARIABLE LABELS filter_$ 'ZRE_2 > 3 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

```
FREQUENCIES VARIABLES=Passieve OPL lft man migratieachtergrond werksituatie
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
/ORDER=ANALYSIS.
```

Actieve

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	72	91,1	91,1	91,1
	1,00	7	8,9	8,9	100,0
Total		79	100,0	100,0	

OPL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	3	3,8	3,8	3,8
	1,00	24	30,4	30,4	34,2
	2,00	52	65,8	65,8	100,0
	Total	79	100,0	100,0	

werksituatie

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	39	49,4	49,4	49,4
	1,00	31	39,2	39,2	88,6
	2,00	9	11,4	11,4	100,0
	Total	79	100,0	100,0	

migratieachtergrond

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	70	88,6	88,6	88,6
	1,00	2	2,5	2,5	91,1
	2,00	7	8,9	8,9	100,0
	Total	79	100,0	100,0	

Statistics

		Actieve	OPL	lft	man	migratieachtergrond	werksituatie
N	Valid	79	79	79	79	79	79
	Missing	0	0	0	0	0	0
Mean		,0886	1,6203	-,3177	,4051	,2025	,6203
Std. Deviation		,28599	,56168	13,90592	,49404	,58572	,68508
Minimum		,00	,00	-26,90	,00	,00	,00
Maximum		1,00	2,00	29,10	1,00	2,00	2,00

DFBETA

Kijken we naar de DFBETA van zelfstandige van model B, dan zijn dat zeven zelfstandigen die mee hebben gedaan aan actieve politieke participatie. De te hoge waarde betekent dat deze respondenten de coëfficiënt omhoog hebben getrokken.

USE ALL.

```

COMPUTE filter_$=(DFB1_2> 0.044).
VARIABLE LABELS filter_$ 'ZRE_1 > 3 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
    
```

```

FREQUENCIES VARIABLES=Actieve OPL lft man migratieachtergrond werksituatie
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
/ORDER=ANALYSIS.
    
```

Actieve

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1,00	9	100,0	100,0	100,0

OPL

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1,00	9	100,0	100,0	100,0

werksituatie

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 2,00	9	100,0	100,0	100,0

migratieachtergrond

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ,00	7	77,8	77,8	77,8
2,00	2	22,2	22,2	100,0
Total	9	100,0	100,0	

man

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ,00	4	44,4	44,4	44,4
1,00	5	55,6	55,6	100,0
Total	9	100,0	100,0	

		lft			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	-27,90	1	11,1	11,1	11,1
	-11,90	1	11,1	11,1	22,2
	-2,90	2	22,2	22,2	44,4
	-,90	1	11,1	11,1	55,6
	5,10	1	11,1	11,1	66,7
	7,10	1	11,1	11,1	77,8
	8,10	1	11,1	11,1	88,9
	16,10	1	11,1	11,1	100,0
	Total	9	100,0	100,0	

Model C

DFBETA

Model 3 kent meer problemen met de DFBETA. Allereerst heeft deze een te lage waarde bij zelfstandige met -0,212, maar daarnaast ook te lage waardes bij eerste generatie (-0,069) en tweede generatie (-0,081). Eerst kijken we naar de respondenten die een te lage DFBETA hebben bij zelfstandige. Dit zijn hoogopgeleide zelfstandigen die niet hebben meegedaan aan passieve politieke participatie. Waarschijnlijk kunnen we deze te lage waarden dus verklaren doordat van deze respondenten wel werd verwacht dat ze passief hadden geparticipeerd, maar dat niet hebben gedaan. De respondenten die een te hoge DFBETA bij eerste generatie hebben zijn dus eerste generatie immigranten met een hoog opleidingsniveau die niet hebben meegedaan aan passieve politieke participatie. Tot slot zijn het bij de te hoge DFBETA van tweede generatie wederom hoogopgeleiden die niet hebben meegedaan. Deze keer zijn het dus immigranten van de tweede generatie.

USE ALL.

```
COMPUTE filter_$=(DFB1_3< -0.056).
VARIABLE LABELS filter_$ 'ZRE_1 > 3 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
```

EXECUTE.

```
FREQUENCIES VARIABLES=Passieve OPL lft man migratieachtergrond werksituatie
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
/ORDER=ANALYSIS.
```

Passieve

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	5	100,0	100,0	100,0

OPL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2,00	5	100,0	100,0	100,0

werksituatie

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2,00	5	100,0	100,0	100,0

migratieachtergrond

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	4	80,0	80,0	80,0
	1,00	1	20,0	20,0	100,0
	Total	5	100,0	100,0	

man

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	2	40,0	40,0	40,0
	1,00	3	60,0	60,0	100,0
	Total	5	100,0	100,0	

lft

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	-18,90	1	20,0	20,0	20,0
	-15,90	1	20,0	20,0	40,0
	-12,90	1	20,0	20,0	60,0
	-8,90	1	20,0	20,0	80,0
	13,10	1	20,0	20,0	100,0
	Total	5	100,0	100,0	

USE ALL.

COMPUTE filter_\$=(DFB3_3<-0.056).

VARIABLE LABELS filter_\$ 'ZRE_1 > 3 (FILTER)'.
 VALUE LABELS filter_\$ 0 'Not Selected' 1 'Selected'.
 FORMATS filter_\$ (f1.0).
 FILTER BY filter_\$.
 EXECUTE.

FREQUENCIES VARIABLES=Passieve OPL lft man migratieachtergrond werksituatie
 /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
 /ORDER=ANALYSIS.

Passieve

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ,00	11	100,0	100,0	100,0

OPL

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 2,00	11	100,0	100,0	100,0

werksituatie

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ,00	5	45,5	45,5	45,5
1,00	5	45,5	45,5	90,9
2,00	1	9,1	9,1	100,0
Total	11	100,0	100,0	

migratieachtergrond

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1,00	11	100,0	100,0	100,0

man

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ,00	4	36,4	36,4	36,4
1,00	7	63,6	63,6	100,0
Total	11	100,0	100,0	

lft

		Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	-23,90	1	9,1	9,1	9,1	
	-19,90	1	9,1	9,1	18,2	
	-17,90	1	9,1	9,1	27,3	
	-16,90	1	9,1	9,1	36,4	
	-14,90	1	9,1	9,1	45,5	
	-11,90	1	9,1	9,1	54,5	
	-9,90	1	9,1	9,1	63,6	
	-2,90	1	9,1	9,1	72,7	
	-,90	1	9,1	9,1	81,8	
	9,10	1	9,1	9,1	90,9	
	13,10	1	9,1	9,1	100,0	
	Total		11	100,0	100,0	

USE ALL.

COMPUTE filter_\$=(DFB4_3< -0.056).

VARIABLE LABELS filter_\$ 'ZRE_1 > 3 (FILTER)'.
 VALUE LABELS filter_\$ 0 'Not Selected' 1 'Selected'.
 FORMATS filter_\$ (f1.0).
 FILTER BY filter_\$.
 EXECUTE.

FREQUENCIES VARIABLES=Passieve OPL lft man migratieachtergrond werksituatie
 /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
 /ORDER=ANALYSIS.

Passieve

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	14	100,0	100,0	100,0

OPL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2,00	14	100,0	100,0	100,0

werksituatie

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	6	42,9	42,9	42,9
	1,00	8	57,1	57,1	100,0
Total		14	100,0	100,0	

migratieachtergrond

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2,00	14	100,0	100,0	100,0

man

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	10	71,4	71,4	71,4
	1,00	4	28,6	28,6	100,0
	Total	14	100,0	100,0	

lft

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	-21,90	2	14,3	14,3	14,3
	-20,90	1	7,1	7,1	21,4
	-19,90	1	7,1	7,1	28,6
	-17,90	4	28,6	28,6	57,1
	-16,90	1	7,1	7,1	64,3
	-14,90	2	14,3	14,3	78,6
	-8,90	1	7,1	7,1	85,7
	-4,90	1	7,1	7,1	92,9
	-2,90	1	7,1	7,1	100,0
	Total	14	100,0	100,0	

Deelconclusie

Modellen A, B en C kennen dus elk wel problemen met problemen met invloedrijke punten. Voor passieve politieke participatie komt dat met name doordat er respondenten zijn voor wie de geschatte kans op politieke participatie vrij hoog is (vaak bij hoogopgeleiden), maar in werkelijkheid niet hebben meegedaan aan passieve politieke participatie. Het gevolg is dat de coëfficiënten van de variabelen met de te lage DFBETA lager zijn dan wanneer deze respondenten niet in de steekproef zouden zitten. Voor model B, waar actieve politieke participatie werd geschat, lijkt er juist sprake te zijn van het tegenovergestelde: van respondenten werd niet verwacht dat ze hadden meegedaan aan politieke participatie, maar hebben dat desondanks wel gedaan.

Ondanks dat deze respondenten een relatief grote invloed uitoefenen op de uiteindelijke resultaten, laten we deze respondenten in de steekproef, want er is geen reden om aan te nemen dat de scores van deze respondenten onjuist of onrealistisch zijn.

Multicollineariteit

Multicollineariteit is afhankelijk van de mate waarin onafhankelijke variabelen met elkaar samenhangen. Wanneer deze te hoog is, kunnen de afzonderlijke variabelen minder variantie van de afhankelijke variabele verklaren, doordat deze ook al wordt verklaard door andere variabelen. Er ontstaan grotere standaardfouten en de afzonderlijke effecten van de variabelen in de regressiemodellen zijn minder sterk. De effecten kunnen daardoor wegvallen of hun significantie verliezen. Ter controle schatten we in plaats van binaire logistische regressiemodellen nu lineaire regressiemodellen waarbij de VIF-scores zijn berekend. Deze score staat voor in hoeverre de verklaarde variantie ook door andere variabelen wordt verklaard. De VIF-scores spiegelen we aan een kritieke waarde van 4. Uiteraard nemen we de interactievariabelen niet mee, gezien die vanzelfsprekend in sterke mate samenhangen met de variabelen waarvan ze een product zijn.

Na het schatten van de regressiemodellen zien we dat er geen problemen zijn van multicollineariteit. Geen van de waarden benadert de kritieke waarde: alle waarden zitten tussen de 1 en de 2. Er is dus weinig sprake van overlap in verklaarde variantie.

REGRESSION

```
/MISSING LISTWISE
/STATISTICS COLLIN TOL
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT Passieve
/METHOD=ENTER OPLlaag OPLhoog Zelfstandige wnpriv Gen1 Gen2 man lft.
```

REGRESSION

```
/MISSING LISTWISE
/STATISTICS COLLIN TOL
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT Actieve
/METHOD=ENTER OPLlaag OPLhoog Zelfstandige wnpriv Gen1 Gen2 man lft.
```

```
COMPUTE filter_$=(OPLhoog = 1 & Coo_1>0).
VARIABLE LABELS filter_$ 'OPLhoog = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

REGRESSION

```
/MISSING LISTWISE
/STATISTICS COLLIN TOL
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT Passieve
/METHOD=ENTER OPLlaag OPLhoog Zelfstandige wnpriv Gen1 Gen2 man lft.
```

REGRESSION

/MISSING LISTWISE
 /STATISTICS COLLIN TOL
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT Actieve
 /METHOD=ENTER OPLlaag OPLhoog Zelfstandige wnpriv Gen1 Gen2 man lft.

Model A

Coefficients^a

Model		Collinearity Statistics	
		Tolerance	VIF
1	OPLlaag	,827	1,209
	OPLhoog	,815	1,226
	Zelfstandige	,883	1,132
	wnpriv	,815	1,227
	Gen1	,986	1,014
	Gen2	,970	1,030
	man	,960	1,042
	lft	,905	1,104

a. Dependent Variable: Passieve

Model B

Coefficients^a

Model		Collinearity Statistics	
		Tolerance	VIF
1	OPLlaag	,827	1,209
	OPLhoog	,815	1,226
	Zelfstandige	,883	1,132
	wnpriv	,815	1,227
	Gen1	,986	1,014
	Gen2	,970	1,030
	man	,960	1,042
	lft	,905	1,104

a. Dependent Variable: Actieve

Model C

Coefficients^a

Model		Collinearity Statistics	
		Tolerance	VIF
1	Zelfstandige	,902	1,109
	wnpriv	,827	1,209
	Gen1	,977	1,024
	Gen2	,971	1,030
	man	,923	1,084
	lft	,894	1,118

a. Dependent Variable: Passieve

Model D

Coefficients^a

Model		Collinearity Statistics	
		Tolerance	VIF
1	Zelfstandige	,902	1,109
	wnpriv	,827	1,209
	Gen1	,977	1,024
	Gen2	,971	1,030
	man	,923	1,084
	lft	,894	1,118

a. Dependent Variable: Actieve

Modelfit

Tot slot kijken we naar de fit van de modellen, dus in hoeverre de modellen goed in staat zijn om de variabele te voorspellen. Daarvoor kijken we naar drie uitkomsten van de analyses: de *deviance*, de *Hosmer and Lemeshow test* en de classificatietabellen. De *deviance* is een maat van in hoeverre het model in *fit* vergelijkbaar is aan het perfecte model, waarbij een hogere score staat voor een slechtere *fit*. Een 0-score betekent een perfect model. Deze maat vergelijken we in elke stap van de modellen. De *Hosmer Lemeshow test* maakt op basis van voorspelde kansen 10 groepen en test in deze groepen in hoeverre de voorspelde proporties overeenkomen met de geobserveerde proporties van de afhankelijke variabele. De test toetst de nulhypothese dat het model een slechte voorspeller van de afhankelijke variabele is. Dus een lagere *p*-waarde betekent een slechter model. Dat bekijken we alleen voor de uiteindelijke modellen. Als laatst kijken we nog naar de classificatie tabellen: deze vergelijkt de geschatte proporties van de afhankelijke variabele met de daadwerkelijke proporties. Door de classificatietabel van het eerste model met het laatste model te vergelijken, kunnen we

beoordelen in hoeverre de modellen betere voorspellers zijn geworden. Deze maten gaan we per model bij langs.

```
LOGISTIC REGRESSION VARIABLES Passieve
/METHOD=ENTER lft man Gen1 Gen2 wnpriv
/METHOD=ENTER Zelfstandige OPLlaag OPLhoog
/METHOD=ENTER OPLlaag*Zelfstandige OPLhoog*Zelfstandige
/PRINT=GOODFIT CI(95)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

Model A: Passieve politieke participatie

Deviance

Voor de *deviance* kijken we naar de waarde van de *-2 Log likelihood en X^2 -toets*. In blok 1 is die waarde 1659,013. Aan de waarde van de R^2 zien we dat dit model nog ver van het perfecte model verwijderd is. Maar, we zien ook aan de X^2 -toets ($X^2=67,173, p<0,001$) dat toevoeging van de controlevariabelen aan het model wel voor significante verbetering van de *fit* zorgt. In blok 2 zijn de onafhankelijke variabelen toegevoegd, waarmee de *deviance* daalt naar 1601,34. Deze daling is niet heel groot, maar wel significant ($X^2=57,658, p<0,001$). Tot slot komen daar in blok 3 de interactievariabelen erbij. We zien dat deze maar heel weinig veranderen aan de *deviance*: 1601,34 ($X^2=0,040, p=0,980$). Toevoeging van de interactievariabelen helpt dus maar heel weinig om beter de afhankelijke variabele te kunnen voorspellen.

Blok 1

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1659,013 ^a	,025	,053

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

Step		Chi-square	df	Sig.
Step 1	Step	67,173	5	<,001
	Block	67,173	5	<,001
	Model	67,173	5	<,001

Blok 2

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1601,354 ^a	,047	,097

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

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	57,658	3	<,001
	Block	57,658	3	<,001
	Model	124,831	8	<,001

Blok 3

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1601,314 ^a	,047	,097

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

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	,040	2	,980
	Block	,040	2	,980
	Model	124,871	10	<,001

Hosmer Lemeshow test

Voor de *Hosmer and Lemeshow test* kijken we alleen naar het uiteindelijke model. Zoals beschreven, deelt deze test de steekproef op in 10 groepen op basis van de voorspelde kans. Vervolgens wordt de nul-hypothese getoetst die stelt dat het model een slechte *fit* heeft. Het gaat daarbij niet per sé om het toetsten van de hypothese, maar meer om het beoordelen van de *p*-waarde.

Wanneer we kijken naar de toets, dan zien we dat deze een zeer lage *p*-waarde heeft ($X^2=20,332$, $p=0,009$), wat dus betekent dat het model een slechte fit heeft. Wanneer we kijken naar de tabel met de proporties, dan zien we er een paar groepen zijn die slecht zijn voorspeld: 4, 5 en 8. Gekeken naar de andere groepen, dan lijkt het model een redelijk goede voorspeller.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	20,332	8	,009

Contingency Table for Hosmer and Lemeshow Test

		Passieve = ,00		Passieve = 1,00		Total
		Observed	Expected	Observed	Expected	
Step 1	1	62	68,488	196	189,512	258
	2	48	42,440	212	217,560	260
	3	33	34,395	228	226,605	261
	4	43	29,819	222	235,181	265
	5	15	24,545	245	235,455	260
	6	16	20,688	244	239,312	260
	7	20	17,222	244	246,778	264
	8	20	13,359	240	246,641	260
	9	5	10,047	256	250,953	261
	10	6	6,996	249	248,004	255

Classificatietabellen

Tot slot kijken we naar de classificatietabellen. Daarbij vergelijken we de classificatietabel van het nul-model (alleen een constante) met die van het uiteindelijke model. Zo kunnen we beoordelen of het uiteindelijke model een betere voorspeller is gebleken.

Helaas geven de classificatietabellen weinig informatie, doordat voor bijna iedereen wordt verwacht dat ze wel meedoen aan passieve politieke participatie. Dat komt doordat de *cut-off value* van de tabellen ligt op 0,500. Gezien het grootste deel van de steekproef wel meedoet aan passieve politieke participatie, is de kans voor bijna alle respondenten groter dan 0,500. We zien daarom maar een zeer kleine verbetering in het percentage correct voorspelde respondenten (van 89,7% naar 89,8%).

Classification Table^{a,b}

Observed		Predicted		Percentage Correct
		Passieve ,00	1,00	
Step 0	Passieve ,00	0	268	,0
	1,00	0	2336	100,0
Overall Percentage				89,7

a. Constant is included in the model.

b. The cut value is ,500

Classification Table^a

	Observed	Predicted		Percentage Correct
		Passieve ,00	1,00	
Step 1	Passieve ,00	2	266	,7
	1,00	0	2336	100,0
Overall Percentage				89,8

a. The cut value is ,500

Deelconclusie

Het uiteindelijke model blijkt dus niet een hele goede voorspeller van de afhankelijke variabele. Vooral de toevoeging van de interactievariabelen zorgt voor nauwelijks betere voorspellingen. Wel is het uiteindelijke model volgens de X^2 -toets een significant betere voorspeller dan het nul-model.

Model B: Actieve politieke participatie

```
LOGISTIC REGRESSION VARIABLES Actieve
/METHOD=ENTER lft man Gen1 Gen2 wnpriv
/METHOD=ENTER Zelfstandige OPLlaag OPLhoog
/METHOD=ENTER OPLlaag*Zelfstandige OPLhoog*Zelfstandige
/PRINT=GOODFIT CI(95)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

Deviance

Allereerst kijken we weer naar de *deviance* van de modellen. In blok 1 met de controlevariabelen is de *deviance*: 2403,83 met $X^2=33,759$ en $p<0,001$. Het model heeft dus een significant betere *fit* dan het nul-model. In blok 2 is die *deviance* gedaald naar 2367,58 met $X^2=36,249$ en $p<0,001$. De *fit* van het model is dus significant verbeterd, maar de *deviance* is niet heel sterk gedaald. In blok 3 zorgt toevoeging van de interactievariabelen niet voor een significante verbetering van de *fit* van het model: *deviance*=2365,9, $X^2=1,665$, $p=0,435$.

Blok 1

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2403,826 ^a	,013	,021

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

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	33,759	5	<,001
	Block	33,759	5	<,001
	Model	33,759	5	<,001

Blok 2

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2367,577 ^a	,027	,044

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

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	36,249	3	<,001
	Block	36,249	3	<,001
	Model	70,008	8	<,001

Blok 3

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2365,912 ^a	,027	,045

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

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	1,665	2	,435
	Block	1,665	2	,435
	Model	71,673	10	<,001

Hosmer and Lemeshow test

Dit model komt veel beter uit de *Hosmer and Lemeshow test* dan het model dat passieve politieke participatie verklaarde. In dit model is de $X^2=7,119$ met $p=0,524$. De p -waarde is veel hoger, wat betekent dat de nulhypothese (het model is een slechte voorspeller) minder waarschijnlijk is.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	7,119	8	,524

Contingency Table for Hosmer and Lemeshow Test

		Actieve = ,00		Actieve = 1,00		Total
		Observed	Expected	Observed	Expected	
Step 1	1	242	239,837	21	23,163	263
	2	233	232,629	28	28,371	261
	3	221	226,066	39	33,934	260
	4	224	220,756	35	38,244	259
	5	211	218,816	49	41,184	260
	6	214	214,322	47	46,678	261
	7	215	206,319	43	51,681	258
	8	205	206,258	58	56,742	263
	9	203	195,697	57	64,303	260
	10	173	180,301	86	78,699	259

Classificatietabellen

Kijken we naar de classificatietabel dan zien we hetzelfde probleem als bij passieve politieke participatie, alleen is nu de geschatte kans in alle gevallen lager dan de *cut-off value* van 0,5. Voor iedereen in de steekproef is de kans op actieve politieke participatie dus lager dan 0,5, wat ervoor zorgt dat voor niemand actieve politieke participatie wordt voorspeld.

Classification Table^{a,b}

	Observed	Predicted		Percentage Correct
		Actieve ,00	Actieve 1,00	
Step 0	Actieve ,00	2141	0	100,0
	Actieve 1,00	463	0	,0
Overall Percentage				82,2

a. Constant is included in the model.

b. The cut value is ,500

Classification Table^a

Observed		Predicted		Percentage Correct
		Actieve ,00	1,00	
Step 1	Actieve ,00	2141	0	100,0
	1,00	463	0	,0
Overall Percentage				82,2

a. The cut value is ,500

Deelconclusie

Net als bij het model van passieve politieke participatie, veranderen toevoegingen van variabelen weinig aan de *deviance* van de modellen al zijn deze veranderingen wel significant. Daarmee zorgen deze variabelen voor significant betere voorspellingen. Dit wordt onderstreept door de uitkomst van de *Hosmer and Lemeshow test*, waarbij de nulhypothese die stelt dat het model een goede *fit* heeft een redelijk hoge *p*-waarde heeft. Wederom zeggen de classificatietabellen weinig over de *fit* van het uiteindelijke model.

Model C: Passieve politieke participatie geselecteerd op hoog opleidingsniveau

```
COMPUTE filter_$=(OPLhoog = 1 & Coo_1>0).
VARIABLE LABELS filter_$ 'OPLhoog = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

```
LOGISTIC REGRESSION VARIABLES Passieve
/METHOD=ENTER lft man Gen1 Gen2 wnpriv
/METHOD=ENTER Zelfstandige
/PRINT=GOODFIT CI(95)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

Deviance

In blok 1 heeft het model een deviance van 527,9 met $X^2=41,368$ en $p<0,001$. Toevoeging van de controlevariabelen zorgen door voor significante verbetering van de *fit* van het model. In blok 2 is zelfstandige toegevoegd als onafhankelijke variabele. Daar zien we dat dit voor een nauwelijks betere *fit* zorgt: *deviance*=527,9 met $X^2=0,039$ en $p=0,844$.

Blok 1

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	527,896 ^a	,032	,089

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

Omnibus Tests of Model Coefficients

Step 1		Chi-square	df	Sig.
	Step	41,368	5	<,001
	Block	41,368	5	<,001
	Model	41,368	5	<,001

Blok 2

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	527,857 ^a	,032	,089

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

Omnibus Tests of Model Coefficients

Step 1		Chi-square	df	Sig.
	Step	,039	1	,844
	Block	,039	1	,844
	Model	41,406	6	<,001

Hosmer and Lemeshow test

Voor deze toets wordt de nulhypothese, het model heeft een slechte *fit*, niet verworpen. Desondanks heeft het model wel een redelijk lage *p*-waarde. Dat zien we terug in de tabel waar opvalt dat het model vooral moeite heeft om mensen die niet meedoen aan passieve politieke participatie goed te voorspellen.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	11,881	8	,157

Contingency Table for Hosmer and Lemeshow Test

		Passieve = ,00		Passieve = 1,00		Total
		Observed	Expected	Observed	Expected	
Step 1	1	24	22,816	102	103,184	126
	2	9	12,011	119	115,989	128
	3	7	8,481	119	117,519	126
	4	4	7,264	124	120,736	128
	5	10	6,094	117	120,906	127
	6	9	5,207	119	122,793	128
	7	5	4,274	122	122,726	127
	8	1	3,588	127	124,412	128
	9	2	2,959	125	124,041	127
	10	4	2,308	116	117,692	120

Classificatietabellen

Net als eerder zorgen de classificatietabellen niet voor veel inzicht in de *fit* van het model: voor geen van de respondenten is de kans op passieve politieke participatie lager dan de *cut-off value* van 0,5.

Classification Table^{a,b}

Observed	Passieve	Predicted		Percentage Correct
		,00	1,00	
Step 0	Passieve ,00	0	75	,0
	1,00	0	1190	100,0
Overall Percentage				94,1

a. Constant is included in the model.

b. The cut value is ,500

Classification Table^a

Observed	Passieve	Predicted		Percentage Correct
		,00	1,00	
Step 1	Passieve ,00	0	75	,0
	1,00	0	1190	100,0
Overall Percentage				94,1

a. The cut value is ,500

Model D: Actieve politieke participatie geselecteerd op hoog opgeleiden.

LOGISTIC REGRESSION VARIABLES Actieve
 /METHOD=ENTER lft man Gen1 Gen2 wnpriv
 /METHOD=ENTER Zelfstandige
 /PRINT=GOODFIT CI(95)
 /CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).

Deviance

De *deviance* van blok 1 met de controlevariabelen is 1326,38 met $X^2=20,974$ en $p<0,001$. Het model kent dus een significante verbetering van de *fit* ten opzichte van het nul-model. In blok 2 heeft toevoeging van zelfstandige maar weinig invloed op de hoogte van de *deviance* 1324,85 met $X^2=1,532$ en $p=0,216$. Het zorgt dus niet voor een significante verbetering van de *fit*.

Blok 1

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1326,383 ^a	,016	,025

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

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	20,974	5	<,001
	Block	20,974	5	<,001
	Model	20,974	5	<,001

Blok 2

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1324,851 ^a	,018	,027

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

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	1,532	1	,216
	Block	1,532	1	,216
	Model	22,506	6	<,001

Hosmer and Lemeshow test

De nulhypothese van de *Hosmer and Lemeshow test* wordt niet verworpen ($X^2=9,368$, $p=0,312$), wat betekent dat het model een redelijke *fit* heeft. Voor de meeste groepen is de het uiteindelijke model een redelijke voorspeller.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	9,368	8	,312

Contingency Table for Hosmer and Lemeshow Test

Step 1		Actieve = ,00		Actieve = 1,00		Total
		Observed	Expected	Observed	Expected	
1	1	111	111,591	19	18,409	130
2	2	108	109,402	23	21,598	131
3	3	96	104,211	32	23,789	128
4	4	108	101,384	19	25,616	127
5	5	98	97,358	26	26,642	124
6	6	100	98,294	27	28,706	127
7	7	100	96,712	27	30,288	127
8	8	98	95,007	30	32,993	128
9	9	91	88,187	33	35,813	124
10	10	71	78,855	48	40,145	119

Classificatietabellen

Tot slot zien we ook bij de classificatietabellen van dit model hetzelfde probleem: geen van de respondenten heeft een hogere kans dan 0,5 op politieke participatie, dus behaalt niemand de *cut-off value*. Hierdoor zien we in de classificatietabellen geen verbetering van het model.

Classification Table^{a,b}

Observed	Actieve	Predicted		Percentage Correct
		,00	1,00	
Step 0	Actieve ,00	981	0	100,0
	1,00	284	0	,0
Overall Percentage				77,5

a. Constant is included in the model.

b. The cut value is ,500

Classification Table^a

Observed	Actieve	Predicted		Percentage Correct
		,00	1,00	
Step 1	Actieve ,00	981	0	100,0
	1,00	284	0	,0
Overall Percentage				77,5

a. The cut value is ,500

Deelconclusie

Het model is een redelijke voorspeller van de afhankelijke variabele, maar toevoeging van de onafhankelijke variabele zelfstandige doet daar weinig aan. Het model heeft desondanks een redelijke *fit*.