

# **Alle scholen zijn gelijk, maar sommige meer dan andere**

Een ruimtelijke analyse van basisschoolresultaten  
in Noord-Nederland.

---

**Bachelorthesis Marten Hoekstra**

**Rijksuniversiteit Groningen**

**Faculteit Ruimtelijke Wetenschappen**

**Begeleider Jessica Vogelzang**

**Januari 2014**

## Samenvatting

De recentelijk gepubliceerde eindtoetsscores van basisscholen in Nederland heeft de mogelijkheid gecreëerd om sociaalgeografisch onderzoek te doen naar de relatie tussen omgevingsomstandigheden en schoolprestaties. Onderwijsprestaties worden beïnvloed door tal van zaken; individuele leerling omstandigheden, familieachtergrond, kwaliteit van individuele scholen en leraren. In dit onderzoek wordt de rol van de sociaaleconomische status (SES) van een postcode gebied op schoolprestaties onderzocht. Bewijs voor dit verband wordt geleverd met behulp van een meervoudige lineaire regressie waarin drie factoren van de SES zijn meegenomen; inkomen, aandeel allochtonen en percentage uitkeringontvangers. Voor het verband tussen inkomen & schoolprestaties en aandeel allochtonen & schoolprestaties wordt bewijs gevonden. Deze verbanden tonen aan dat er een substantiële relatie is tussen de sociaaleconomische omstandigheden en basisschool prestaties.

## Voorwoord

Geachte lezer,

Bij het kiezen van een onderwerp voor mijn bachelorthesis kon ik twee thema's bij elkaar brengen die mijn bijzondere belangstelling hebben; geografie en onderwijs. Toen ten tijde van het bedenken van een thema ook de discussie los barstte over de CITO toets resultaten heb ik niet lang meer getwijfeld over dit thema als onderwerp voor mijn bachelor scriptie. Maar een thema is één ding, het uitwerken daarvan een tweede (en derde). Daarbij heb ik van veel mensen hulp gehad die ik graag wil bedanken.

Mijn eerste dank gaat uit naar mijn begeleider bij de Faculteit Ruimtelijke Wetenschappen: Jessica Vogelzang. Bij het kiezen van een onderzoeksgebied, het uiteindelijke thema en het vormgeven van het onderzoeksvoorstel heeft haar kritische en praktische blik geholpen bij het vlot en gedegen doorlopen van het eerste deel van het bachelorproject. Jessica's professionele feedback tijdens het (oneindig) analyseren van de data en het schrijven van deze thesis heeft mij geholpen om kritisch te kijken naar de thesis. Maar bovenal was de begeleiding erg plezierig en succesvol.

Daarnaast gaat mijn dank uit naar Mark Verlaat van het CIT voor het beschikbaar stellen van de topografische kaarten die zijn gebruikt om de kaarten van deze thesis te creëren. Ook wil ik alle medewerkers en docenten van de faculteit bedanken die mij konden helpen om korte en iets uitgebreidere vragen te beantwoorden.

Ten slotte wil ik studiegenoten, vrienden en familie bedanken voor hun tips, opmerkingen, gedachtewisselingen en feedback. Hun 'externe' blik heeft mij enorm geholpen tijdens het gehele proces. In het bijzonder wil ik mijn broer, Hylke, bedanken voor zijn scherpe en zeer bruikbare commentaren.

Rest mij u, de lezer en geïnteresseerde, veel leesplezier te wensen. Voor vragen en opmerkingen houd ik mij graag aanbevolen.

Marten Hoekstra

Groningen, januari 2014

*hoekstra.marten@live.nl*

# Inhoudsopgave

1. <b>Inleiding</b> .....	1
1.1 Aanleiding.....	1
1.2 Probleemstelling.....	2
1.3 Opbouw van de thesis .....	2
2. <b>Theoretisch kader</b> .....	3
2.1 Ontwikkeling onderwijs binnen sociale geografie.....	3
2.2 Nederlands onderzoek .....	4
2.3 Conclusie theoretisch kader .....	5
2.4 Conceptueel model .....	6
3. <b>Data en methodologie</b> .....	7
3.1 Data onderwijsresultaten .....	7
3.2 Data sociaaleconomische omstandigheden.....	10
3.3 Methodologie .....	13
4. <b>Resultaten</b> .....	15
4.1 Model met drie variabelen.....	15
4.2 Model met twee variabelen .....	17
5. <b>Discussie</b> .....	18
6. <b>Conclusie</b> .....	20
7. <b>Literatuurlijst</b> .....	22
8. <b>Bijlagen</b> .....	24
9.1 Bijlage I .....	24
9.2 Bijlage II .....	25
9.3 Bijlage III .....	26

# 1. Inleiding

## 1.1 Aanleiding

Het vergelijken van basisscholen op basis van hun prestaties is lange tijd niet (goed) mogelijk geweest. Lang heeft het primair onderwijs zich afhoudend gehouden tegenover vergelijkend onderzoek op basis van resultaten (Sykes & Musterd, 2010). Enerzijds vanwege het ontbreken van bruikbare data en valide onderzoeksmethoden en anderzijds vanwege de angst bij basisscholen en koepelorganisaties voor het publiceren van kwaliteitscijfers; schoolbesturen in Nederland houden de boot af vanwege mogelijke negatieve gevolgen bij het vergelijken van schoolresultaten door ouders. (Beter Onderwijs Nederland, 2013)

Vanaf 2015 zijn basisscholen verplicht om aan het einde van de schoolcarrière van leerlingen een Eindtoets Basisonderwijs af te nemen (Rijksoverheid, 2013a). Een bekende en veelgebruikte toets in het primair onderwijs is de CITO toets. Naast de CITO eindtoets, zijn er ook andere varianten. Met deze eindtoets willen basisscholen samen met eerdere behaalde resultaten leerlingen een advies geven voor vervolgonderwijs. Deze toets is in de eerste plaats bedoeld voor individuele leerlingen maar wordt ook gebruikt als instrument voor de onderwijsinspectie. Vanuit de maatschappij luidt steeds vaker de vraag om deze resultaten publiekelijk te maken om zodoende het brede publiek te informeren over onderwijsprestaties (Nederlands Dagblad, 2013).

In september 2013 kreeg RTL Nieuws van de rechter toestemming om alsnog de eindtoets resultaten van Nederlandse basisscholen te publiceren (PO-raad, 2013). Het wereldkundig maken van de resultaten kwam RTL Nieuws op veel kritiek te staan; vanuit de onderwijswereld had men bezwaar op het vrijgeven van de resultaten zonder rekening te houden met lokale en sociale omstandigheden (Wij-leren.nl, 2013). Schoolbestuurders en onderwijsorganisaties beargumenteren dat het openbaar maken van zogenoemde 'ruwe' eindtoets resultaten zonder aanvullende context geen recht doet aan individuele scholen met specifieke lokale omstandigheden (Beter Onderwijs Nederland, 2013).

Juist die redenen; de nieuwe openbare bron van data (CITO resultaten) en de kritiek van het onderwijsveld (het ontbreken van de link tussen lokale omstandigheden en eindtoetsuitslagen), waren dan ook aanleiding om, onder andere op basis van de verkregen data, een wetenschappelijke ruimtelijke analyse uit te voeren en het primair onderwijs geografisch te onderzoeken. En daarmee onderzoek, dat in het buitenland al eerder en grondiger is onderzocht, op de Nederlandse context te projecteren en zodoende een bijdrage te leveren aan de ruimtelijke analyse van basisschoolprestaties.

## 1.2 Probleemstelling

Het maatschappelijke vraagstuk rond de transparantie van basisschool resultaten zijn de aanleiding om een geografisch onderzoek te doen naar de relatie tussen schoolprestaties in de vorm van eindtoets scores en omgevingsfactoren. Met omgevingsfactoren worden de sociaaleconomische omstandigheden bedoeld (zoals inkomen, percentage werklozen, sociale samenstelling buurt en wijk). De hoofdvraag die in dit onderzoek wordt beantwoord is als volgt:

*In hoeverre is er een relatie tussen basisschool prestaties en lokale sociaaleconomische factoren in Noord-Nederland?*

Daarop aansluitend zijn er de volgende deelvragen;

1. Welke verbanden zijn er tussen resultaten van basisscholen en het inkomen in Noord-Nederland en hoe sterk zijn deze verbanden?
2. Welke verbanden zijn er tussen resultaten van basisscholen en het percentage allochtonen in Noord-Nederland en hoe sterk zijn deze verbanden?
3. Welke verbanden zijn er tussen werk(loosheid) en de resultaten van basisscholen en hoe sterk zijn deze verbanden?

## 1.3 Opbouw van de thesis

Deze thesis is als volgt opgebouwd; allereerst wordt het onderzoek in het *Theoretisch kader* in een breder perspectief geplaatst en de resultaten van recente vergelijkbare onderzoeken geanalyseerd en met elkaar vergeleken. De conclusies van eerdere en vergelijkbare onderzoeken vormen de basis voor nieuwe hypotheses die in deze thesis worden getoetst. In de volgende sectie, *Data en methodologie*, worden de keuzes die zijn gemaakt bij het verzamelen van de data, het al dan niet bewerken van de gegevens en het kiezen van de methodologie verantwoord. De resultaten van de kwantitatieve data-analyse worden in de paragraaf *Resultaten* gepresenteerd. De resultaten worden in de paragraaf *Discussie* verder toegelicht en mogelijke verbanden aangetoond. De thesis wordt afgesloten met de conclusie waarin de onderzoeksvragen met behulp van de resultaten worden beantwoord.

## 2. Theoretisch kader

### 2.1 Ontwikkeling onderwijs binnen sociale geografie

Het onderzoeksveld 'onderwijs' binnen de sociale geografie is betrekkelijk nieuw. Andere (semi)publieke sectoren, zoals gezondheidszorg en publiek vastgoed hebben lange tijd de aandacht gekregen van sociaalgeografen (Hamnett & Butler, 2011). Vanaf de laatste eeuwwisseling is er sprake van een toename van onderzoek naar de geografie van onderwijs (Hanson Thiem, 2009). In de minder dan twintig jaar dat er uitgebreid internationaal onderzoek naar is gedaan, is het onderzoeksveld uitgegroeid tot een significant onderdeel binnen de geografie (Sykes & Kuyper, 2009; Holloway et al., 2010). Enkele recente belangrijke onderzoeken waren Hamnett & Butler (2011), Butler & Hamnett, (2010), Walker en Clark, (2010) en Singleton et al.(2011). Binnen de geografie van onderwijs is vooral de afgelopen tien jaar een ander terrein ontdekt; de relatie tussen omgeving/de lokale (sociaaleconomische) omstandigheden en de resultaten van scholen en/of leerlingen individueel (Ainsworth, 2002; Hanson Thiem, 2009). Ander onderzoek van Gibbons & Machin (2006) weet (deels) een link te leggen tussen schoolprestaties en huizenprijzen.

Onderzoek van Gordon & Monastiriotis (2007) en Webber & Butler (2007) proberen met GCSE resultaten (nationale toets voor middelbare scholieren in het Verenigd Koninkrijk) een ruimtelijke analyse te maken van schoolresultaten. Gordon & Monastiriotis (2007) doen dit door leerlingen op basis van een woonplaats te classificeren en daarna de schoolprestaties te verbinden met deze buurten met hogere of lagere sociaaleconomische statussen. In het Verenigd Koninkrijk, met relatief veel segregatie op lokaal niveau wellicht een goede insteek, maar voor onderzoek naar de Nederlandse omgevingen is de aanpak van Webber & Butler (2007) geschikter; schoolresultaten worden statistisch getoetst aan een keur van variabelen (percentage alleenstaande ouders, allochtonen, moedertaal, opleiding, werk, etc.). Het onderzoek van McCulloch (2006) gebruikt andere data (cognitieve vaardigheidstesten bij kinderen van 4-9 jaar) maar past wel in het rijtje met soortgelijke onderzoeken. McCulloch (2006) onderzocht het verband tussen typen buurten met hogere en lagere sociaaleconomische status (SES) en resultaten op cognitieve- en gedragsvaardigheden. Voor buurten met een lage SES werd een significant verband aangetoond tussen lokale omstandigheden en resultaten op desbetreffende testen.

Dat de bovengenoemde onderzoeken veelal plaats hebben gevonden in Groot-Brittannië is geen toeval; het Britse onderwijssysteem kent traditiegetrouw veel onderscheid tussen maatschappelijke klassen en mede daardoor vermeende kwaliteitsverschillen (Butler & Hamnett, 2007; Martin & Atkinson, 2001). Dit gecombineerd met een eind jaren '80 geïntroduceerde nieuwe neoliberale agenda waarin scholen niet langer een product van de maatschappij zijn, maar meer markt gedreven

instanties zijn waarin voor de gebruikers (ouders en leerlingen) wat te kiezen valt, wekte de interesse van geografen en beleidsmakers voor een ruimtelijk onderzoek naar onderwijsresultaten (Gulson & Symes, 2007). Butler & Hamnett (2007) hebben de conclusies van verscheidene Britse onderzoeken naast elkaar gelegd en stellen dat de kwaliteit van de leefomgeving terug te zien is in de prestaties van scholen.

Naast het Britse onderzoek is er ook in Zweden onderzoek gedaan op dit gebied. In onderzoek van Andersson & Subramanian (2006) wordt gesteld dat *“buurteigenschappen gecombineerd met sociaaleconomische middelen en demografische stabiliteit uitstekende voorspellers zijn van individuele schoolprestaties.”* (p.2022). Onderzoek van Brännström (2008), eveneens uitgevoerd in Zweden, wijst uit dat sommige buurteigenschappen meer invloed hebben op schoolprestaties, maar dat de lokale school het meest bepalend is voor de uitkomsten in leerprestaties; scholen vlakken negatieve buurteigenschappen af, als een soort ‘mediator’. In de Verenigde Staten is door Ainsworth (2002) onderzocht wat de invloed is van de buurt op Amerikaanse middelbare school prestaties en concludeert dat de economische positie van families van leerlingen een grote rol spelen in het voorspellen van lokale schoolprestaties.

In Helsinki, Finland, heeft Kauppinen (2007) met data van ruim tienduizend jongeren onderzoek gedaan naar het effect van buurtomstandigheden op de schoolprestaties van diezelfde leerlingen. De specifieke buurt eigenschappen zijn in dit onderzoek opgesteld met behulp van het percentage hoge inkomens, opleidingsniveau en het aandeel werkende inwoners. In dit onderzoek is er bewijs gevonden om de relatie tussen de leefomgeving en het niveau van onderwijs aan te tonen. Vervolgonderzoek van Kauppinen (2008) bevestigt het onderzoek van Brännström (2008) dat individuele scholen het effect van de buurt verkleint, maar dat er wel degelijk nog een verband is tussen de lokale sociaaleconomische status (SES) en onderwijs resultaten. Zo komen leerlingen met een hogere SES significant vaker in een preacademische opleiding.

## **2.2 Nederlands onderzoek**

Soortgelijke onderzoeken zijn in mindere mate in Nederland uitgevoerd. Vaak richten zich deze op etniciteit, segregatie en de maatschappelijke positie van migranten (Karsten et al., 2006). Sykes & Kuypers (2009) en Sykes & Musterd (2010) hebben met andere beschikbare data dan de eindtoets scores een grote bijdrage geleverd aan het onderzoek naar de relatie tussen buurt/omgeving en de schoolprestaties van middelbare scholieren. De belangrijkste conclusies uit Sykes & Kuypers (2009): Leerlingen met een hoge sociaaleconomische status worden minder beïnvloed door de omgeving en niet alle leerlingen zijn vertegenwoordigd in het onderzoek, kleine groepen worden verdrongen in de statistieken. Het onderzoek van Sykes & Musterd (2010) komt met een tegenovergestelde conclusie;



*“the associations between neighbourhood characteristics and achievement were reduced to non-significance.”* (Sykes & Musterd 2010, p.1). Sykes & Kuypers (2009) stellen dat onderwijsresultaten te voorspellen zijn uit twee soorten variabelen: individuele- en buurtvariabelen. Individuele variabelen zijn hier leeftijd, geslacht, afkomst en Familie SES. Buurtvariabelen zijn percentage mensen met hoog/laag inkomen, werkloosheidsuitkeringen, gemiddeld inkomen, aandeel huurwoningen, percentage etnische minderheden, urbanisatiegraad, aantal inwoners omgeving/dorp/stad. Mede vanwege het feit dat dit onderzoek betreft in een Nederlandse setting, met vergelijkbare data die voor dit onderzoek ook beschikbaar zijn is het onderzoek van Sykes & Kuypers (2009) bijzonder waardevol.

Het onderzoek van Webber & Butler (2007) en Gordon & Monastiriotis (2007) hebben, weliswaar in een andere context, een soortgelijke basis van onderzoek; het voorspellen van resultaten op basis van woonplaats respectievelijk een ruimtelijke analyse van de voortgezet onderwijs resultaten. Eén van de belangrijkste conclusie die Webber & Butler (2007) trekken is dat omgevingsfactoren minstens zo belangrijk zijn als persoonlijke en sociale omstandigheden. Maar in hetzelfde onderzoek is een andere conclusie dat *“de buurten die het slechtste scoren op onderwijsgebied zijn niet de buurten die bovengemiddelde achterstand hebben (sociaaleconomisch red.)”* (Webber & Butler 2007, p. 1252.) In het onderzoek van Butler & Hamnett (2007) is een conclusie; *“Some of the best performances (...) come from some ethnic minorities.”* (Butler & Hamnett 2007, p.1170). Daar sluit een volgende bevinding van Butler & Hamnett (2007) op aan; een mix van positieve en negatieve sociaaleconomische omstandigheden kan juist een significant versterkend effect hebben op onderwijsresultaten.

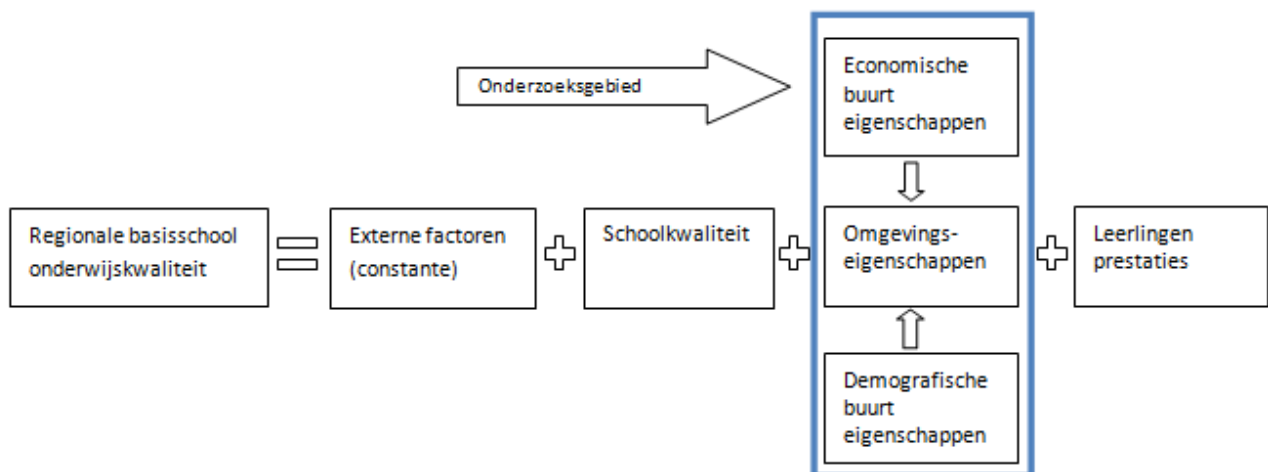
### **2.3 Conclusie theoretisch kader**

Hoewel de conclusies van de verschillende onderzoeken niet altijd met elkaar rijmen, of soms zelfs conflicteren, is er in de wetenschappelijke publicaties consensus dat lokale omgeving- en maatschappelijke factoren een belangrijke rol spelen in onderwijsprestaties. (Gordon & Monastiriotis, 2007; Sykes & Kuyper, 2008; Sykes & Musterd, 2011; Östh et al., 2012). Of zoals Webber & Butler (2006) treffend formuleren:

*“Whatever the reason, there is no doubt that neighbourhood is an effective predictor of most behaviours. (...) It is clear that, whilst school leadership, ethos and teacher motivation can be of importance, none of these has anything like the importance of social composition.”* (Webber & Butler 2006, p.1243)

Veel onderzoek, zoals bovenstaande Nederlandse van Sykes & Kuyper (2009) en Sykes & Musterd (2010), richt zich in het bijzonder op scholieren vanaf twaalf jaar, de resultaten van scholieren in het Nederlandse primaire onderwijs is tot op heden niet of nauwelijks geografisch onderzocht. Mede vanwege de argumenten aangevoerd door de diverse primair onderwijsorganisaties (het meenemen van lokale omstandigheden en niet alleen focussen op de ‘ruwe’ cijfers) en de conclusies van de hier genoemde onderzoeken is het daarom interessant om aanvullend onderzoek uit te voeren. Gezien de sterke conclusies van bijvoorbeeld Butler & Hamnett (2007) en McCulloch (2006) is het interessant om te kijken of deze bevindingen ook toe te passen zijn in de Nederlandse context: Zijn er ook hier verbanden tussen omgeving en schoolprestaties en zijn deze net zo sterk als in de genoemde onderzoeken?

## 2.4 Conceptueel model



Om het onderzoek en de plaats van het onderzoek te verduidelijken is er een conceptueel model samengesteld. Zie figuur 1 voor het model.

**Figuur 1. Conceptueel model onderzoek ruimtelijke analyse basisschoolprestaties (bron: eigen bewerking)**

Het theoretisch kader veronderstelt dat de regionale basisschool onderwijskwaliteit wordt beïnvloed door constante externe factoren, individuele schoolkwaliteiten, omgevingseigenschappen en individuele leerling prestaties. De invloed van de omgevingseigenschappen op de regionale onderwijskwaliteit is het gebied waar deze thesis onderzoek naar doet. De omgevingseigenschappen zijn in dit onderzoek opgesplitst in economische- en demografische buurteigenschappen.

### 3. Data en methodologie

Om de hoofdvraag, *In hoeverre is er een relatie tussen primaire onderwijsresultaten en regionale sociaaleconomische factoren in Noord-Nederland?*, te beantwoorden is er een kwantitatieve analyse uitgevoerd met enerzijds onderwijsdata (eindtoetsresultaten van basisscholen) en anderzijds sociaaleconomische data (economische en demografische buurteigenschappen). Het hoge aantal cases (zie ook 3.1 Data onderwijsresultaten) sluit een kwalitatieve analyse vanwege praktische redenen uit. Alle data voor dit onderzoek zijn samengesteld uit, of herleidt tot het zogenaamde postcode-4 niveau. In Noord-Nederland zijn dat de postcodes 7740 tot en met 7999 en 8350 tot en met 9999. Een postcode-4 gebied bevat gemiddeld 1400 huishoudens.

#### 3.1 Data onderwijsresultaten

De gegevens die zijn gebruikt om de onderwijsresultaten van scholen en daarmee indirect ook de omgeving (het verband tussen scholen en omgeving wordt hierna besproken) te bepalen zijn afkomstig van het Ministerie van Onderwijs, Cultuur en Wetenschap (hierna ministerie van onderwijs genoemd), en bewerkt door RTL Nieuws. Van de 7.112 basisscholen (inclusief speciaal onderwijs) die in Nederland zijn gevestigd (CBS, 2013) zijn ruim zevenduizend scholen in dit overzicht vertegenwoordigd. Naast de eindtoets scores van 2010, 2011 en 2012, zijn er ook gegevens over de vestigingsplaats, denominatie, visie school, bestuursvorm en afkomst leerlingen in opgenomen. Omdat dit onderzoek zich richt op de relatie tussen basisschool prestaties en de sociaaleconomische omstandigheden ter plekke, is van dit bestand alleen gebruik gemaakt van de eindtoetsscores als factor voor de basisschool prestaties.

Voor dit onderzoek zijn 1054 Drentse, Friese en Groningse scholen geselecteerd. In deze selectie zijn ook de basisscholen voor speciaal onderwijs te vinden zoals scholen voor moeilijk lerende kinderen of voor leerlingen met een fysieke of mentale beperking. Kenmerkend voor deze scholen is dat leerlingen relatief verder weg wonen van hun school dan bij leerlingen op een reguliere basisschool. Dit heeft gevolgen voor de uit te voeren analyse en de bijbehorende conclusies; voor correcte gegevens moeten de gegevens extra bewerkt worden en kunnen er uit de statistische toetsen minder sterke conclusies worden getrokken (Sykes & Musterd, 2010). Uitgangspunt bij dit onderzoek is, omwille van de beperking van complexiteit in het databestand, om zoveel mogelijk gebruik te maken van oorspronkelijke data.

Voor de eindtoets, die sinds maart 2013 verplicht is gesteld (Rijksoverheid, 2013), hebben scholen keuze uit Cito Eindtoets Basisonderwijs (beter bekend als de Cito-toets), drempelonderzoek (of drempeltoets), Entreetoets, Leerlingvolgsysteem (LVS), Didactische Leeftijds Equivalenten (DLE).

Daarnaast zijn er enkele scholen die andere, eigen toetsen afnemen. Veruit de meeste basisscholen, 863, maken gebruik van Cito-toets. Dit onderzoek heeft zich gericht op deze 863 scholen. Gekozen is om de andere toetsvormen niet mee laten doen, de belangrijkste reden hiervoor is dezelfde als de keuze om speciaal basisonderwijs uit te sluiten voor dit onderzoek; de beschikbare data niet of zo weinig mogelijk bewerken. Wanneer deze gegevens wel zouden worden toegevoegd waren één of meer bewerkingen onvermijdelijk en dit zou de betrouwbaarheid van conclusies niet ten goede komen (Sykes & Kuyper, 2009).

In het databestand van het ministerie van onderwijs zijn de gemiddelde Cito scores van de desbetreffende scholen weergegeven. Dat wil zeggen; de gemiddelde Cito score (minimum score 500, maximum score 550) behaalt door de leerlingen die deelnemen aan de toets. Om hoge of lage toevalstreffers bij bijvoorbeeld kleine scholen met weinig deelnemende leerlingen te elimineren is er gekozen om de behaalde resultaten per school nogmaals te middelen, waarbij er minimaal twee toets jaren bekend moeten zijn. Omdat niet bij elke school er ten minste twee dezelfde toetsen afgenomen zijn, of de uitslagen daarvan bekend zijn, vielen er 101 scholen af en is het scholenbestand gekrompen tot 762 scholen.

Ten slotte dient er nog één schakel te worden gelegd tussen de resultaten van basisscholen en de postcodes waar leerlingen wonen. Een groot deel van de ouders kiest de dichtstbijzijnde school, maar voor bijvoorbeeld specifieke denominatie (gereformeerd, katholiek) of visie (Jenaplan, Montessori) zullen leerlingen niet altijd in hun eigen omgeving terecht kunnen. Kort gezegd; de locatie van een school is niet gelijk aan de woonplaats van leerlingen. Leerlingen wonen vaak, maar niet per se in de buurt van hun school. Het is dan ook niet mogelijk om de resultaten van scholen direct te koppelen aan de plaats waarin de school gevestigd is.

Om toch de schoolresultaten te koppelen aan postcodes waar de leerlingen wonen die naar desbetreffende school gaan is er gebruik gemaakt van gegevens verstrekt door de Dienst Uitvoering Onderwijs (DUO), die met behulp van Basisregister Onderwijs (BRON) en de Basisregister Instellingen (BRIN) per 4-cijferig postcode het aantal leerlingen per schoolvestiging heeft verzameld. Bijvoorbeeld voor Gemeente Aa en Hunze is bekend dat de drie leerlingen van postcode 9511 naar één basisschool gaan. Op deze manier kan de ruimtelijke spreiding van leerlingen worden bepaald, en met deze stap is het dan ook verantwoord om de basisschool prestaties te koppelen aan postcodes.

Voor de meeste postcodes geldt dat leerlingen naar meerdere scholen gaan. Door middel van een gewogen gemiddelde is er voor deze postcodes ook een prestatiecijfer (Cito-toets gemiddelde) bekend. Dit wordt toegelicht in het volgende voorbeeld; voor postcode 8635 zijn er in totaal 8

leerlingen die naar twee scholen gaan. De eerste school ontvangt drie leerlingen, de tweede vijf. Ze hebben respectievelijk een prestatiecijfer van 532,7 en 538,3. Door het prestatiecijfer van de eerste school voor  $3/8^e$  mee te tellen en het tweede voor  $5/8^e$  kan het gemiddelde prestatiecijfer voor postcode 8635 worden berekend: 536,2. Op basis van deze methode en de 762 scholen is er voor 940 postcode-4 gebieden een gemiddeld prestatiecijfer berekend.

Om deze procedure te verduidelijken zijn in figuur 2 en 3 een kleine, fictieve dataset opgenomen van de twee bronnen van gegevens; Cito-toets uitslagen en woonplaats leerlingen.

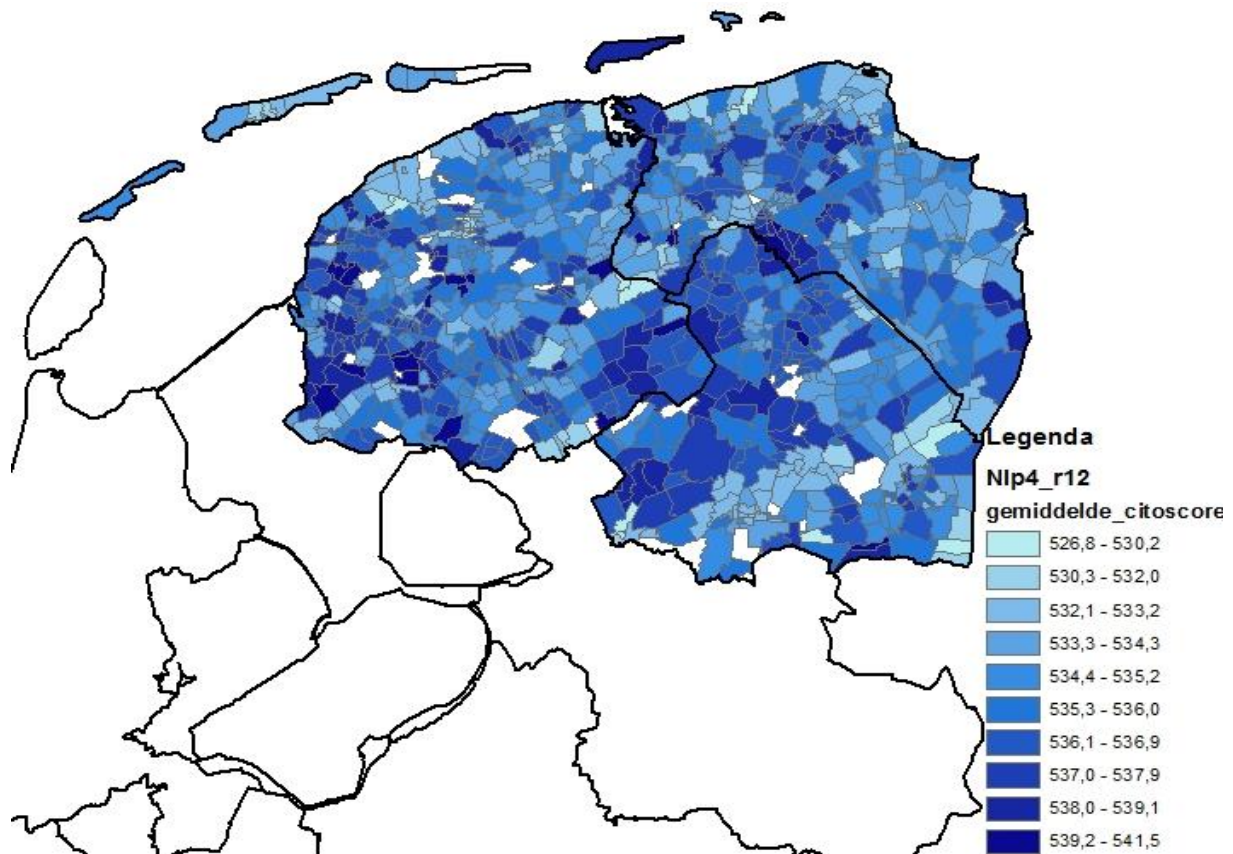
Basisschool	Gemiddelde CITO uitslag
Basisschool A	534,6
Basisschool B	532,3

**Figuur 2. Voorbeeld van CITO uitslagen (bron: Ministerie van Onderwijs, bewerkt).**

Postcode leerlingen	Basisschool	Aantal leerlingen in deze postcode
1110	Basisschool A	8
1111	Basisschool A	3
1111	Basisschool B	4
1112	Basisschool B	9

**Figuur 3. Voorbeeld van DUO woonplaats leerlingen (bron: DUO, bewerkt).**

Om per postcode het gemiddelde prestatiecijfer te berekenen moeten de gegevens van figuur 2 en 3 worden gecombineerd. In dit voorbeeld wonen op postcode 1110 acht leerlingen die naar één basisschool gaan, dus de gemiddelde CITO score is 534,6. Voor postcode 1111 wordt de berekening als volgt:  $(3/7^e \times 534,6) + (4/7^e \times 532,3) = 533,3$ . Postcode 1112 is uiteraard 532,2. In figuur 4 zijn de resultaten van de 940 postcodes in een kaart weergegeven. De gemiddelde scores lopen van 526,8 tot en met 541,5. De witte gebieden in de drie noordelijke provincies geven aan dat hier om eerder genoemde redenen geen resultaten van bekend zijn.



Figuur 4. Gemiddelde CITO score per postcode-4 gebied. (Bron: eigen bewerking)

### 3.2 Data sociaaleconomische omstandigheden

Om de sociaaleconomische omstandigheden te bepalen kan er theoretisch uit een breed aanbod factoren worden gekozen die ook in voorgaande onderzoeken aan bod kwamen; urbanisatiegraad, werkloosheid, type woning, aandeel allochtonen, opleidingsniveau, inkomen, uitkering. In de praktijk zijn er op postcode-4 niveau slechts een beperkt aantal gegevens beschikbaar. De gekozen indicatoren voor de sociaaleconomische status zullen hier worden besproken.

Zoals eerder in het *Theoretisch kader* is beschreven, is voor de factor sociaaleconomische omstandigheden één indicator van doorslaggevend belang: (bruto) inkomen (Webber & Butler, 2007, Hanson Thiem, 2008). De meeste inkomens data in Nederland worden samengesteld op basis van een regionaal, provinciaal of gemeentelijk schaalniveau. Gegevens over inkomen op lokaal niveau, zoals postcode-4, is slechts beperkt beschikbaar en wordt in tegenstelling tot de grotere schaalniveaus niet gepubliceerd en/of jaarlijks vernieuwd.

Het Centrum voor Beleidsstatistiek van het Centraal Bureau voor de Statistiek (CBS-CvB) heeft in 2008 zijn meest recente versie van het *Gemiddeld persoonlijk en besteedbare (bruto)*

*huishoudinkomen naar postcodegebied* gepubliceerd. Aan het Regionaal Inkomensonderzoek (RIO) 2008 is op persoonsniveau de postcode-4 (PC4) gekoppeld van het meest recente adres in 2008. Op basis daarvan is per postcodegebied het gemiddelde persoonlijke en het gemiddelde besteedbare huishoudinkomen berekend. Bij de berekening van het gemiddelde persoonlijke en besteedbare huishoudinkomen zijn per PC4 alleen particuliere huishoudens met een inkomen in 2008 meegenomen. Studentenuishoudens, institutionele huishoudens (zoals bejaardentehuizen) en particuliere huishoudens zonder inkomen zijn buiten dit onderzoek gelaten.

In dit onderzoek is het gemiddelde besteedbare huishoudinkomen afgerond op honderdtallen. Het aantal personen en huishoudens is afgerond op tientallen. Wanneer het aantal huishoudens per postcodegebied kleiner is dan 70 of het aantal personen kleiner dan 200, zijn de gegevens niet bekend gemaakt. Onder besteedbaar huishoudinkomen wordt het volgende verstaan; het bruto-huishoudinkomen verminderd met betaalde inkomensoverdrachten, premies inkomensverzekeringen, premies ziektekosten verzekeringen en belastingen op inkomen en vermogen. Van de 1172 noordelijke postcodes zijn er in dit overzicht 741 met een bekend gemiddeld besteedbaar huishoudinkomen.

Een terechte vraag bij de inkomensgegevens is of de gegevens uit 2008 gebruikt kunnen worden met bijvoorbeeld de schoolprestaties uit 2010 en later. Deze opmerking kan met twee argumenten worden weerlegd; inkomenseffect werkt vertraagd door naar omgevingseffecten, en dus ook naar prestaties van basisscholen. Een hoger of lager inkomen zal op de korte termijn geen direct effect hebben, pas op de langere termijn (meerdere jaren) heeft het uitwerking. Een tweede rechtvaardiging om de data wel te gebruiken is de aanname dat inkomens weliswaar jaarlijks veranderen, maar geografische verschillen zullen op korte termijn niet worden weggewerkt; de verhouding tussen het besteedbare inkomen in de regio Winschoten en de stad Groningen zal over een periode van twee á drie jaar geen substantiële veranderingen laten zien. Bovendien is er, gezien het hoge aantal cases (741), geen reden om de ruimtelijke spreiding van het besteedbare huishoudinkomen in twijfel te trekken.

Naast inkomen wordt ook de sociale compositie van de omgeving gezien als factor om de sociaaleconomische status (SES) van desbetreffende buurt te bepalen (Walker & Clark, 2010; Sykes & Kuyper, 2009). In Nederland publiceert het CBS elk jaar een monitor *Bevolking en Huishoudens* met op lokaal niveau (buurt en postcode) gegevens over bevolkingssamenstelling, waaronder het percentage (niet-) westerse allochtonen. Sykes & Kuyper (2009) maken in hun onderzoek gebruik van deze variabele als factor voor de bepaling van de SES. De meest recente gegevens van het CBS op postcode-4 niveau zijn afkomstig uit 2012. 983 noordelijke postcodes zijn in dit bestand opgenomen.

De volgende gegevens zijn afkomstig uit dit bestand: Totaal aantal inwoners, totaal aantal allochtone inwoners, aantal westerse allochtonen, aantal niet-westerse allochtonen. Gegevens zijn absolute getallen, voor gebruik in statistische toetsen zijn deze omgezet in relatieve cijfers (aantal (niet-)westerse) allochtonen/totaal aantal inwoners x 100%). Alle cases hebben op deze wijze een percentage variërend van nul tot 73% (totaal allochtonen), nul tot 28% (westerse-) en nul tot 57% (niet-westerse allochtonen) gekregen.

Ten slotte is er een derde factor die in eerdere onderzoeken relevant bleek te zijn. Zowel Webber & Butler (2007) als Sykes & Kuyper (2009) als Andersson (2004) als Andersson & Subramanian (2006) nemen in hun onderzoeken het aandeel werklozen c.q. uitkeringsgerechtigden/uitkeringsontvangers mee in hun model. Data voor, in dit geval, uitkeringsontvangers is verzameld door het Bureau voor Beleidsstatistiek van het CBS voor het jaar 2009. In opdracht van Sinfore heeft het Bureau voor Beleidsstatistiek een inventarisatie gemaakt voor huisartsen en zorginstellingen zodat zij toekomstig beleid kunnen aanpassen aan lokale sociaaleconomische omstandigheden. In deze zogenaamde *wijkscan* is ook het aantal uitkeringsontvangers opgenomen. Helaas is deze data niet compleet; van veel postcodes zijn de gegevens überhaupt niet bekend en bij andere postcodes zijn de inwoneraantallen laag en is het privacy gevoelige informatie. In het onderzoek van het CBS zijn 876 noordelijke postcodes opgenomen, daarvan zijn 103 postcodes met informatie over werkloosheidsuitkeringen (11,8% van het totale databestand). Rekening houdende met het relatief hoge aantal cases (103) is besloten om deze factor wel mee te laten wegen.

De kaarten met de ruimtelijke gegevens van de afzonderlijke onderscheiden factoren (huishoudeninkomen, aandeel niet-westerse allochtonen en percentage uitkeringsontvangers) zijn als bijlage opgenomen in deze thesis. Zie bijlage I, II en III. De postcodes met ontbrekende waarden zijn wit gemarkeerd.

In dit onderzoek zijn de postcodes de te onderzoeken cases. Het gemiddelde prestatiecijfer de afhankelijke variabele; te verklaren uit het percentage allochtonen, het gemiddeld te besteden huishoudeninkomen en het aandeel werkloosheidsuitkeringen. Door de cases uit de eerder genoemde databestanden met elkaar samen te voegen is er voor dit onderzoek een nieuw bestand gecreëerd. Van 919 cases/postcodes zijn het gemiddeld CITO score en percentage allochtonen bekend, van 713 cases is bovendien het gemiddelde huishoudeninkomen bekend. Ter indicatie; in geheel Drenthe, Fryslân en Groningen zijn 979 postcodes.



### 3.3 Methodologie

In dit onderzoek wordt gezocht naar een eventueel verband tussen onderwijsresultaten en de drietal onderscheiden indicatoren voor sociaaleconomische omstandigheden. Met de beschikbare data kan dit worden vertaald naar een statistisch model; een meervoudig regressieanalyse. In een dergelijke analyse oefenen meerdere onafhankelijke variabelen, in dit onderzoek aandeel allochtonen, gemiddeld huishoudeninkomen en aandeel uitkeringsontvangers, invloed uit op de afhankelijke variabele (prestatiecijfer basisschool). Deze statistische methode stelt als voorwaarde dat er sprake is van één afhankelijke variabele met één of meer onafhankelijke interval/ratiovariabelen. Deze onderzoeksmethode in lijn met gebruikte toetsen in vergelijkbare onderzoeken van Sykes & Musterd (2010) en Gordon & Monastiriotis (2006). Andere onderzoeksmethoden zoals Chi-kwadraat toets en factoranalyse vallen af omdat de hier onderscheiden factoren niet los van elkaar kunnen worden gezien; alleen een meervoudige lineaire regressie kan deze type data statistisch toetsen.

Voor dit onderzoek zijn alle cases in een nieuw bestand geplaatst die met behulp van IBM SPSS Statistics Data Editor (SPSS) statistisch zijn onderzocht. Voordat een meervoudige regressieanalyse is uitgevoerd zijn de cases onderzocht op multicollineariteit. Daarnaast is het eventuele verband tussen de drie onafhankelijke variabelen onderzocht op lineariteit, aangezien dit een voorwaarde is voor meervoudige lineaire regressie.

Multicollineariteit, het effect als meerdere onafhankelijke variabelen sterk met elkaar correleren, is niet wenselijk bij meervoudige regressie. Door SPSS de tolerantie en de VIF-waarde te laten berekenen kan er een uitspraak worden gedaan over de al dan niet aanwezige multicollineariteit. In dit geval is er sprake van een tolerantiewaarden tussen de 0,748 en 0,935 en VIF waarden van 1,070 tot 1,337. Hoe kleiner de tolerantie, hoe groter het lineaire verband tussen de variabelen en dus lagere toegevoegde waarde in het model (Norušis, 2010). Een waarde dichtbij één is een indicator voor een laag lineair verband. Hoge VIF waarden, boven de 2,5, gecombineerd met een lage tolerantie wijst op instabiliteit. Daarvan is hier geen sprake, multicollineariteit is niet aanwezig in dit model.

De resultaten van een meervoudige regressie worden sterker naarmate er een sterker lineair verband aanwezig is. De verschillende variabelen zijn in een scatterplot matrix visueel weergegeven om lineaire verbanden aan te tonen. De matrix liet voor schoolprestaties en inkomen een lineair verband zien, voor schoolprestaties en werkloosheidsuitkeringen en percentage allochtonen een minder sterk lineair verband. Om het verband sterker te maken is besloten de data op verschillende manieren te bewerken (Norušis, 2010). Van de variabelen huishoudeninkomen en de schoolprestaties zijn de natuurlijke logaritmes berekend. Deze bewerking heeft het verband niet

verbeterd en is dus niet in het uiteindelijke model geplaatst. Het inkomen en het percentage niet-westerse allochtonen zijn gekwadrateerd, maar deze aanpassing heeft evenmin het beoogde doel bereikt. Het bewerken van de data heeft niet geleid tot verbeteringen van de lineairiteit, om die reden is ervoor gekozen om de data in haar oorspronkelijke vorm te gebruiken en daarmee een meervoudige lineaire regressie uit te voeren.

## 4. Resultaten

In dit onderdeel zullen de uitkomsten van de meervoudige lineaire regressie worden getoond. Daarna worden in hoofdstuk 5 de resultaten van de regressie besproken en verklaard.

De 919 postcodes zijn de cases in het databestand. Van 919 cases is één variabele bekend (percentage niet-westerse allochtonen), van 713 twee variabelen (+ gemiddelde huishoudeninkomen) en van 103 zijn alle drie variabelen bekend (+ werkloosheidsuitkeringen). Bij het uitvoeren van de meervoudige lineaire regressie met alle drie variabelen worden alleen de postcodes gebruikt waarvan de waarden van alle variabelen bekend zijn, in dit geval 103. Omdat er met deze methode veel data verloren gaat, 816 cases worden niet in het model geplaatst, is er gekozen om een tweede soortgelijk model samen te stellen met twee variabelen (gemiddelde huishoudeninkomen en percentage niet-westerse allochtonen) en 713 cases. De tweede regressie is vooral bedoeld als controle model, om de statistische relevantie van beide toetsen te onderstrepen.

### 4.1 Model met drie variabelen

De CITO-scores lopen van 526,8 tot en met 541,5 met als gemiddelde 535,3 en 2,15 als standaardafwijking. Het besteedbaar huishoudeninkomen (18.300-62.100) heeft als gemiddelde 34.167 en een standaardafwijking van 6.096. De percentages niet-westerse allochtonen variëren van 0% tot 56,7% (gemiddeld 2,1%, standaardafwijking 3,9%) en percentages uitkeringen van 19% tot 82% (gemiddeld 40,3%)

Bij het uitvoeren van een lineaire regressie is de validiteit van de statistische toets van belang: Zijn de resultaten van de regressie valide, en mogen ze worden gebruikt om een antwoord te geven op de onderzoeksvragen. Daartoe is bij elke toegevoegde variabele het meervoudige correlatie coëfficiënt ( $R$ ), het determinatiecoëfficiënt ( $R^2$ ), en de *Adjusted R<sup>2</sup>* bekend. Zie figuur 5.

<b>Regressie</b>	0,000
<i>Significantie</i>	
<b>R</b>	0,545
<i>Meervoudige correlatie coefficient</i>	
<b>R<sup>2</sup></b>	0,297
<i>Determinatie coefficient</i>	
<b>Adjusted R square</b>	0,276

**Figuur 5. Meervoudige lineaire regressie (model 1) (1)**

De regressie is significant ( $< 0,0005$ ). Dat wil zeggen dat het model een significant verband heeft aangetoond tussen de afhankelijke en onafhankelijke variabele, dus mag dit model worden gebruikt in dit onderzoek. Een  $R$  neigend naar één geeft een perfecte lineaire regressie aan, nul geen lineair verband. De waarde in dit model, 0,545, past in het beeld van een min of meer lineair verband. Daarbij sluit de  $R^2$  van 0,297 aan: 29,7% van de variantie in het model kan worden verklaard door middel van de variabelen. Met andere woorden, aan de hand van het inkomen, aandeel allochtonen en het percentage uitkeringen kan 29,7% van de cases worden voorspeld. Vanwege de meerdere verklarende variabelen is er in deze toets ook de *Adjusted R square* opgenomen. Deze compenseert voor de variabelen, ook als deze een lage toegevoegde waarde voor het model hebben. In dit geval is de *Adjusted R square* 0,276; 27,6% van de variantie kan worden verklaard met dit model. Alle bovenstaande waarden geven aan dat het model in z'n geheel betrouwbaar en bruikbaar is (Norušis (2010)). Met behulp van figuur 6 kijken we naar de individuele factoren van het model.

	<b>B</b>	<b>t</b>	<b>significantie</b>
<b>Constante</b>	28,859	15,735	0,000
<b>Huishoudeninkomen</b>	0,000	3,467	0,001
<b>Niet-westerse allochtonen</b>	-0,092	-2,404	0,018
<b>Uitkeringontvangers</b>	0,030	1,483	0,141

**Figuur 6. Meervoudige lineaire regressie (model 1) (2)**

Het eerste onderdeel wat van belang is in figuur 6 zijn de significantie cijfers. Waarden onder de 0,05 geven aan dat desbetreffende factor een significante bijdrage aan het model levert. Zoals te zien in figuur 5 passen de variabelen *huishoudeninkomen* en *niet-westerse allochtonen* in het model, de factor uitkeringontvangers heeft geen significante toevoeging. Daarnaast zijn de  $B$  waarden interessant; deze coëfficiënten geven de richting en de sterkte van de richting aan in het model. Ter illustratie; per procentpunt meer niet-westerse allochtonen daalt het gemiddelde CITO-cijfer met 0,092. Het  $B$  coëfficiënt voor huishoudeninkomen is kleiner dan 0,0005, voor niet-westerse allochtonen -0,092 en uitkeringsontvangers 0,030. Maar deze laatste is niet significant en dus niet mee te nemen in de resultaten en conclusie.

## 4.2 Model met twee variabelen

Het tweede model bevat twee verklarende factoren; het aandeel uitkeringontvangers is weggelaten. Daardoor wordt het aantal cases fors uitgebreid (713 in plaats van 103) en bijkomend voordeel is dat, zoals aangetoond in het eerste model, de niet-significante bijdrage van de factor *uitkeringontvangers* is weggelaten. Hierdoor kan het nieuwe model mogelijk sterkere verbanden aantonen. Voor het overige is het model gelaten zoals het was. In figuur 7 en 8 worden de uitkomsten van de nieuwe regressie getoond.

<b>Regressie</b>	0,000
<i>Significantie</i>	
<b>R</b>	0,299
<i>Meervoudige correlatie coefficient</i>	
<b>R<sup>2</sup></b>	0,089
<i>Determinatie coefficient</i>	
<b>Adjusted R square</b>	0,087

**Figuur 7. Meervoudige lineaire regressie (model 2) (1)**

	<b>B</b>	<b>t</b>	<b>significantie</b>
<b>Constante</b>	33,380	68,472	0,000
<b>Huishoudeninkomen</b>	0,00006	4,595	0,000
<b>Niet-westerse allochtonen</b>	-0,108	-4,923	0,000

**Figuur 8. Meervoudige lineaire regressie (model 2) (2)**

De tweede regressie is nog steeds significant ( $< 0,05$ ) maar, zoals verwacht als er een factor buiten de regressie wordt gehouden, is de voorspellende waarde van het model gedaald. Respectievelijk  $R$  en  $R^2$  zijn gezakt tot 0,299 (model 1: 0,545) en 0,089 (model 1: 0,297). De individuele factoren zijn zeer significant; beiden zijn kleiner dan 0,0005. De  $B$  coëfficiënten zijn iets sterker dan in het eerste model en geven daardoor meer richting in het model. Dit is logisch aangezien minder variabelen meer moeten verklaren in deze regressie.

Het beeld wat de uitkomsten van beide regressies laat zien stemt met elkaar overeen; de CITO toets resultaten zijn voor een kleiner of groter deel te verklaren uit de sociaaleconomische factoren. In het volgende deel *Discussie* worden de resultaten verder uitgewerkt en in het licht van eerdere, vergelijkbare onderzoeken geplaatst.

## 5. Discussie

De in dit onderzoek onderscheiden factoren voor sociaaleconomische omstandigheden verklaren voor 27,6% de variantie in het eerste model. Ondanks de verschillen in data worden in het onderzoek van Sykes & Kuyper (2008) exact hetzelfde percentage toegeschreven als het aandeel buurtinvloed op onderwijsresultaten. Iets ouder onderzoek van Harris & Mercier (2000) naar de invloed van een breed scala aan sociaal-culturele en -economische factoren op middelbare schoolprestaties komt op een verklaarde variantie in drie Canadese steden variërend van 0,15 tot 0,36. Gordon & Monastiriotis (2007) hebben een  $R^2$  van 0,25 tot 0,72 met 17 onderscheiden verklarende variabelen, Sykes & Musterd (2010) komen tot 0,28 tot 0,40. Zie figuur 9 voor een schematisch overzicht van de verklaarde varianties in eerdere onderzoeken. De variantie in dit onderzoek ligt dus in lijn met de andere onderzoeken, waar meer factoren zijn meegenomen in de meervoudige lineaire regressie. Met de 'slechts' drie gekozen variabelen in dit onderzoek kan een relatief groot deel van de resultaten worden verklaard. N.b.: Deze lijst is bedoeld om de resultaten van dit onderzoek een plaats te geven in het onderzoeksveld. Het vergelijken van varianties is niet mogelijk vanwege de sterk contextgebonden studies.

Onderzoek (jaar)	Verklaarde variantie ( $R^2$ )
Harris & Mercier (2000)	0,15-0,36
Gordon & Monastiriotis (2007)	0,25-0,72
Sykes & Kuyper (2008)	0,276
Sykes & Musterd (2010)	0,28-0,40
Hoekstra (2014)	0,276

**Figuur 9. Verklaarde variantie van eerdere onderzoeken**

Waar de resultaten van dit onderzoek zich duidelijker afwijken van bijvoorbeeld de resultaten van Harris & Mercier (2000) en Gordon & Monastiriotis (2007) zijn de  $B$  coëfficiënten, oftewel de helling van de lineaire regressie. Deze zijn vergeleken met hellingen in de bovenstaande onderzoeken betrekkelijk laag, 0,000 voor huishoudeninkomen en -0,092 voor niet-westerse allochtonen, in het tweede model is de  $B$  coëfficiënt voor huishoudeninkomen 0,00006. Een getal dat in eerste instantie verwaarloosbaar lijkt maar wat in dit model betekent dat per euro die er gemiddeld in een postcode meer verdiend wordt, de CITO score 0,00006 punten stijgt. Wanneer alle factoren in twee verschillende postcodes gelijk zijn maar het inkomen bijvoorbeeld 25.000 euro hoger is, is de bijbehorende gemiddelde CITO score 1,5 punt hoger in de rijkere postcode. Niet zo sterk als in de Canadese setting van Harris & Mercier (2000), maar zeker in lijn met voorgaand Nederlands

onderzoek van Sykes & Musterd (2010). De nog sterkere relatie tussen inkomen en onderwijsresultaten is grotendeels te verklaren uit de meer gesegeerde buurten en wijken in de onderzochte Canadese steden.

Het *B* coëfficiënt van het aandeel niet-westerse allochtonen in beide regressies ligt rond de waarden die ook in het onderzoek van Harris & Mercier (2000) en Gordon & Monastiriotis (2007) worden gegeven; circa -0,10. Daarbij moet worden aangetekend dat Harris & Mercier (2000) aandeel migranten als factor gebruikt en Gordon & Monastiriotis (2007) etnische minderheden. De Nederlandse onderzoeken van Sykes & Musterd (2010) en Sykes & Kuyper (2008) gebruiken voor etnische achtergronden andere methodes en definities, derhalve kunnen deze uitkomsten niet worden vergeleken.

Ten slotte de niet-significante bijdrage van de factor uitkeringontvangers. In tegenstelling tot de eerder genoemde literatuur van Webber & Butler (2007), Sykes & Kuyper (2009) Andersson (2004) en Andersson & Subramanian (2006), is er in deze regressie geen bewijs voor een significante bijdrage van werkloosheid op de schoolresultaten. Dit is te verklaren aan de hand van in ieder geval twee redenen; het gebruik van andere definities en de gebruikte data in dit onderzoek. Andersson (2004) gebruikt als factor *beroepsstatus* om de sociaaleconomische omstandigheden (deels) te duiden en gebruikt deze gegevens die direct gekoppeld zijn aan de onderzochte cases, in plaats van de gekoppelde databestanden in dit onderzoek. Andersson & Subramanian (2006) maakten in hun onderzoek geen onderscheid tussen werkloosheidsuitkeringen en bijstandsuitkeringen. Van groter invloed is de gebruikte data in dit onderzoek. Van alle cases zijn 103 mét gegevens over de werkloosheidsuitkeringen. Vanwege privacyredenen is veel data (88,9%) niet gebruikt en bovendien zijn de overgebleven cases niet normaal verdeeld. In het databestand is er sprake van een oververtegenwoordiging van cases met hoge aantallen werkloosheidsuitkeringen. Dit heeft er tot geleid dat deze variabele niet het verwachte resultaat heeft laten zien.

## 6. Conclusie

Theoretisch en intuïtief is het zeer aannemelijk dat de sociale en sociaaleconomische context waarin kinderen opgroeien impact heeft op hoe zij zullen presteren op school. In deze studie is onderzocht of er een relatie is tussen de onderwijsprestaties en de lokale sociaaleconomische omstandigheden. Ontelbaar veel factoren spelen mee in de behaalde onderwijsresultaten maar dit onderzoek heeft geprobeerd alle externe factoren buiten beschouwing te laten en op sociaalgeografische wijze het onderwijs ruimtelijk te onderzoeken. Dit onderzoek zal niet concluderen waar de goede en de minder goed presterende scholen zijn gevestigd, juist vanwege de kritiek vanuit de onderwijswereld. Noch zal het welvarende en arme buurten typeren. Een alles verklarend en omvattend model zal het evenmin opleveren.

Wél een conclusie van deze studie; de resultaten van de onderzochte postcodes zijn voor bijna 30 procent te verklaren uit de regressie. Het model is betrekkelijk ongecompliceerd; drie verklarende factoren. Desondanks, of juist dankzij zijn eenvoud is de meervoudige regressie zeer effectief in het voorspellen van waardes. De uitkomsten van het model zijn vergeleken met de resultaten van de meest vooraanstaande en soortgelijke literatuur; Webber & Butler (2007), Sykes & Kuyper (2009) Andersson (2004), Andersson & Subramanian (2006), Harris & Mercier (2000) en Gordon & Monastiriotis (2007).

De relatie tussen basisschool prestaties van een postcode en het bijbehorende inkomen is met deze regressie aangetoond. De helling van deze variabele laat zien dat een hoger inkomen gepaard gaat met wezenlijk hogere gemiddelde CITO uitslagen. Voor het aandeel niet-westerse allochtonen geldt van hetzelfde laken een pak; de significante bijdrage samen het *B* coëfficiënt toont aan dat hoger percentage niet-westerse allochtonen leidt tot lagere gemiddelde CITO uitslagen. Sykes & Kuyper (2009) komen voor Nederlandse middelbare scholieren tot dezelfde conclusie. Het verband tussen onderwijsprestaties en werk wordt, in tegenstelling tot de conclusies van onder andere Webber & Butler (2007), Sykes & Kuyper (2009) en Andersson & Subramanian (2006), niet aangetoond. Daarvoor zijn twee mogelijke verklaringen; er is daadwerkelijk geen verband, of aannemelijker vanwege de literatuur: Er is wel een verband maar deze is niet aangetoond met dit model. Het niet-volledige databestand voor werkloosheidsuitkeringen zal hier waarschijnlijk debet aan zijn.

De bevindingen van Martin & Atkinson (2001) en Gordon & Monastiriotis (2007) in het Verenigd Koninkrijk, van Kauppinen (2008) in Finland, Andersson & Subramanian (2006) in Zweden, Ainsworth (2002) in de Verenigde Staten en de uitkomsten van de Nederlandse onderzoeken van Sykes & Kuyper (2009) en Sykes & Musterd (2010) zijn in één lijn door te trekken naar de Nederlandse basisschool context. Dit betekent dat het model an sich degelijk en betrouwbaar is; de resultaten



wijken niet uitzonderlijk af van vergelijkbare onderzoeken in diverse andere landen en anderzijds geeft juist dat laatste aan dat de situatie in Noord-Nederland identiek is aan plekken met andere sociale en sociaaleconomische omstandigheden. In (Noord) Nederland kan het model worden gebruikt om te bepalen of regio's onder of boven hun verwachte resultaten presteren, rekening houdende met de specifieke lokale omstandigheden.

Door het hoge aantal cases met bijbehorende data is de basis van dit onderzoek sterk gebleken. Desalniettemin zijn er verbeteringen mogelijk in vervolg onderzoek; door databestanden beter op elkaar te laten aansluiten gaat er minder data verloren. Voor de variabele werk(loosheid) is het aanbevelenswaard om te kiezen voor een andere bron van data. Een toekomstig model met volledige data zal wellicht geen andere uitkomsten laten zien, maar is wel sterker en betrouwbaarder. Ten slotte is de bron voor onderwijsdata, CITO gemiddeldes, mogelijk te eenzijdig; het is een momentopname en niet alle scholen en leerlingen doen mee aan deze toets. Een alternatief zou zijn om een uitgebreider leerlingvolgsysteem te hanteren.

Dit onderzoek is uitgevoerd op postcode-4 niveau. Dit biedt zowel mogelijkheden als beperkingen. Het heeft als voordeel dat het onderzoek zich zeer op de lokale situatie richt. Bovendien zijn er veel cases en dus veel resultaten die in de regressie zijn opgenomen. Keerzijde van PC4 is de beperkte hoeveelheid beschikbare data. Het CBS publiceert de meeste lokale informatie op gemeenteniveau en ook andere gegevensverstrekkers werken veelal niet met PC4. Een nieuwe regressie met, desnoods op een kleiner schaalniveau, meer sociaaleconomische variabelen zou het aandeel verklaarde variantie kunnen verhogen en daarmee de voorspellende waarde vergroten.

Met dit onderzoek is een poging gedaan om zowel een gedegen verklaring als voorspelling te geven over de relatie tussen onderwijs en buurt. Met het aantonen van die relatie is er ook erkenning van de directe rol van de omgeving op onderwijsprestaties. Beleidsmakers zullen zich niet alleen moeten richten op (hoge) cijfers, maar op de verwachte resultaten inclusief de invloed van de omgeving daarop.

Ten slotte zijn er nog twee aanbevelingen voor vervolgonderzoek. In dit onderzoek is de rol van de buurt op de onderwijsprestaties onderzocht. Uiteraard is dit geen eenrichtingsverkeer, maar een wisselwerking. Een interessante vervolgstudie zou zijn om te kijken naar wat (verbetering van) onderwijsresultaten doet voor de sociaaleconomische omstandigheden op de lange termijn. Een ander mogelijke studie bouwt voort op resultaten uit dit onderzoek: een kwalitatief onderzoek naar waarom sommige postcodes hoger scoren dan verwacht mag worden aan de hand van de sociaaleconomische omstandigheden.

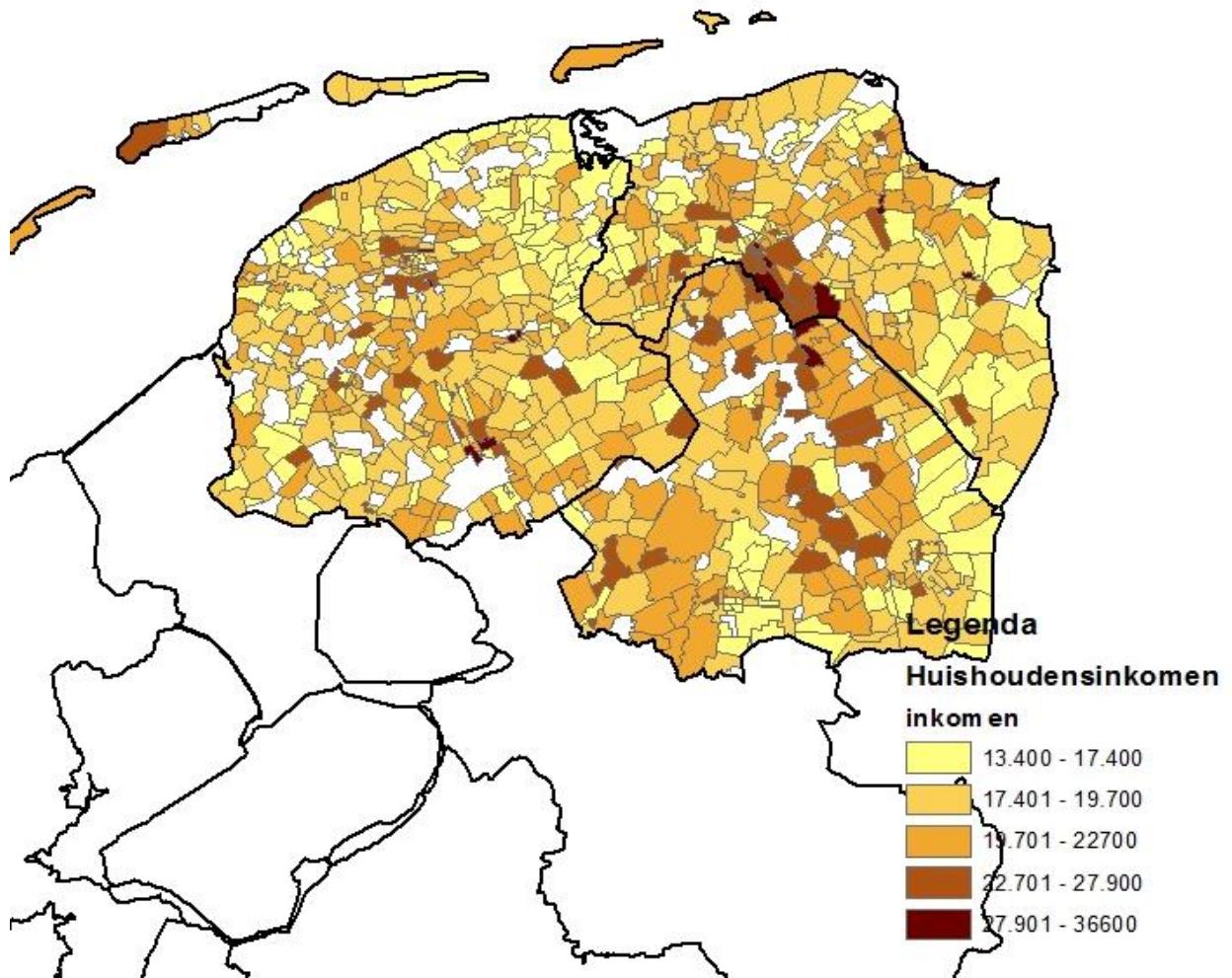
## 7. Literatuurlijst

- Ainsworth, J.W. (2002).** Why Does It Take a Village? The Mediation of Neighborhood Effects on Educational Achievement. *Social Forces*, 81(1), 117-152.
- Andersson, E. (2004).** From valley of sadness to hill of happiness: the significance of surroundings for socioeconomic career. *Urban Studies*, 41, 641-659.
- Andersson, E. & Subramanian, S.V. (2006).** Explorations of Neighbourhood and Educational Outcomes for Young Swedes. *Urban Studies*, 43(11), 2013-2025.
- Bell, C. (2009).** Geography in Parental Choice. *American Journal of Education*, 115(4), 493-521.
- Beter Onderwijs Nederland (2013).** RTL publiceert cito-scores per school. Geraadpleegd op 24-10-2013 via <http://www.beteronderwijsnederland.nl/nieuws/rtl-publiceert-cito-scores-school>
- Brännström, L. (2008).** Making Their Mark: The Effects of Neighbourhood and Upper Secondary School on Educational Achievement. *European Sociological Review*, 24(4), 463-478.
- Butler, T. & Hamnett, C. (2007).** The Geography of Education: Introduction. *Urban Studies*, 44(7), 1161-1174.
- Butler, T. & Hamnett, C. (2010).** 'You take what you are given'; the limits to parental choice in education in east London. *Environment and Planning*, 42, 2431-2450
- Centraal Bureau voor de Statistiek (2013).** CBS Statline Onderwijsinstellingen; grootte en soort. Geraadpleegd op 7-10-2013 via <http://statline.cbs.nl/StatWeb/selection/default.aspx?DM=SLNL&PA=03753&VW=T>
- CITO (2013).** De onzin van rangordelijstjes op basis van de Eindtoets. Geraadpleegd op 20-12-2013 via [http://www.cito.nl/nl/Over%20Cito/pers/onzin\\_rangordelijstjes\\_basis\\_eindtoets.aspx](http://www.cito.nl/nl/Over%20Cito/pers/onzin_rangordelijstjes_basis_eindtoets.aspx)
- Gibbons, S. & Machin, S. (2006).** Paying for primary schools: admission constraints, school popularity or congestion? *The Economic Journal*, 116, C77-C92
- Gordon, I. & Monastiriotis, V. (2007).** Education, Location, Education: A Spatial Analysis of English Secondary School Public Examination Results. *Urban Studies*, 44(7).
- Gulson, K.N. & Symes, C. (2007).** Knowing one's place – Educational theory, policy and the spatial turn in *Spatial Theories of Education* (pp. 1-16). New York; Routledge
- Hamnett, C. & Butler, T. (2011).** 'Geography matters': the role distance plays in reproducing educational inequality in East London. *Transactions of the Institute of British Geographers*, 36, 479-500
- Hanson Thiem, C. (2009).** Thinking through education: the geographies of contemporary educational restructuring. *Progress in Human Geography*, 33(2), 154-173.
- Harris, R. & Mercier, M. (2000).** A test for geographers: the geography of educational achievement in Toronto and Hamilton, 1997. *The Canadian Geographer/Le Géographe canadien*, 3, 210-217.
- Holloway, S.L., Hubbard, P., Jöns, H. & Pimlott-Wilson, H. (2010).** Geographies of education and the significance of children, youth and families. *Progress in Human Geography*, 34(5), 583-600.
- Karels, M. (2013).** Rankinglijst Cito-scores eindtoets verwerpelijk. Geraadpleegd op 24-10-2013 via <http://www.wij-leren.nl/publicatie-cito-resultaten.php>
- Kauppinen, T.M. (2007).** Neighborhood effects in an European city; secondary education of young people in Helsinki. *Social Science Research*. 36, 421-444
- Kauppinen, T.M. (2008).** Schools as mediators of neighbourhood effects on choice between vocational and academic tracks of secondary education in Helsinki. *European Sociological Review*, 24, 379-391.
- Karsten, S., Felix, C., Ledoux, G., Meijnen, W., Roeleveld, J., Van Schooten, E. (2006).** Choosing segregation or integration? The extent and effects of ethnic segregation in Dutch cities, *Education and Urban Society*, 38(2), 228-247.
- Martin, D. & Atkinson, P. (2001).** Investigating the Spatial Linkage of Primary School Performance and Catchment Characteristics. *Geographical & Environmental Modeling*, 5(1), 67-83.

- McCulloch, A. (2006).** Variation in children's cognitive and behavioural adjustment between different types of place in the British National Child Development Study. *Social Science Medicine*, 62, 1865-1879
- Norušis, M.J. (2010).** PASW Statistics 18 Guide to Data Analysis. Prentice Hall inc., Upper Saddle River, New Jersey.
- PO-Raad (2013).** *Beslissing op bezwaar*. Geraadpleegd op 20-12-2013 via [http://www.poraad.nl/sites/www.poraad.nl/files/gebruiker%3A24/beslissing\\_op\\_bezwaar.pdf](http://www.poraad.nl/sites/www.poraad.nl/files/gebruiker%3A24/beslissing_op_bezwaar.pdf)
- Östh, J., Andersson, E. & Malmberg, B. (2012).** School Choice and Increasing Performance Difference: A Counterfactual Approach. *Urban Studies*, 50(2), 407-425.
- Rijksoverheid (2013a).** *Is de citotoets of een andere eindtoets verplicht?* Geraadpleegd op 12-11-2013 via <http://www.rijksoverheid.nl/onderwerpen/basisonderwijs/vraag-en-antwoord/is-de-citotoets-of-een-andere-eindtoets-verplicht.html>
- Rijksoverheid (2013b).** *Wat is de Citotoets?* Geraadpleegd op 16-10-2013 via <http://www.rijksoverheid.nl/onderwerpen/basisonderwijs/vraag-en-antwoord/wat-is-de-citotoets-oftewel-eindtoets-basisonderwijs.html>
- Singleton, A.D., Longley, P.A., Allen, R. & O'Brien, O. (2011).** Estimating secondary school catchment areas and the spatial equity of access. *Computers, Environment and Urban Systems*, 35(2011), 241-249
- Soest, A. van (2013).** Scholen niet in beroep tegen publicatie Cito-scores. *Nederlands Dagblad*, 10-09-2013
- Sykes, B. & Kuyper, H. (2009).** Neighbourhood effects on youth educational achievement in the Netherland: can effects be identified and do they vary by student background characteristics? *Environment and Planning A*, 41(10), 2417-2436.
- Sykes, B. & Musterd, S. (2010).** Examining Neighbourhood and School Effect Simultaneously: What Does the Dutch Evidence Show? *Urban Studies*, 48(7), 1307-1331.
- Walker, M. & Clark, G. (2010).** Parental choice and the primary school: Lifestyle, locality and loyalty. *Journal of Rural Studies*, 26(2010), 241-249.
- Webber, R. & Butler, T. (2007).** Classifying Pupils by Where They Live: How Well Does This Predict Variations in Their GCSE Results? *Urban Studies*, 44(7), 1229-1253.

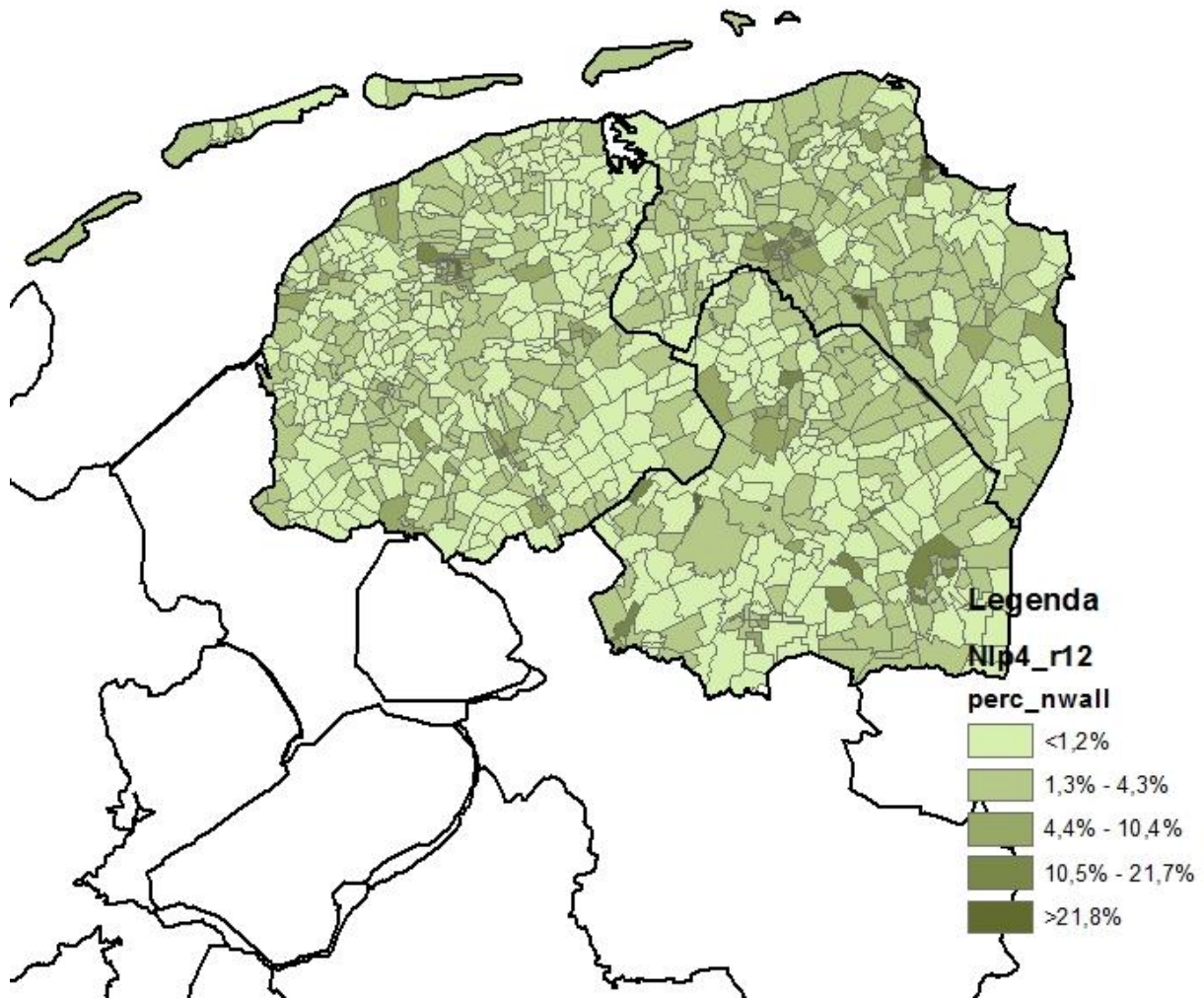
## 8. Bijlagen

### 9.1 Bijlage I



(Bruto)huishoudeninkomen Noord-Nederland (bron: eigen bewerking)

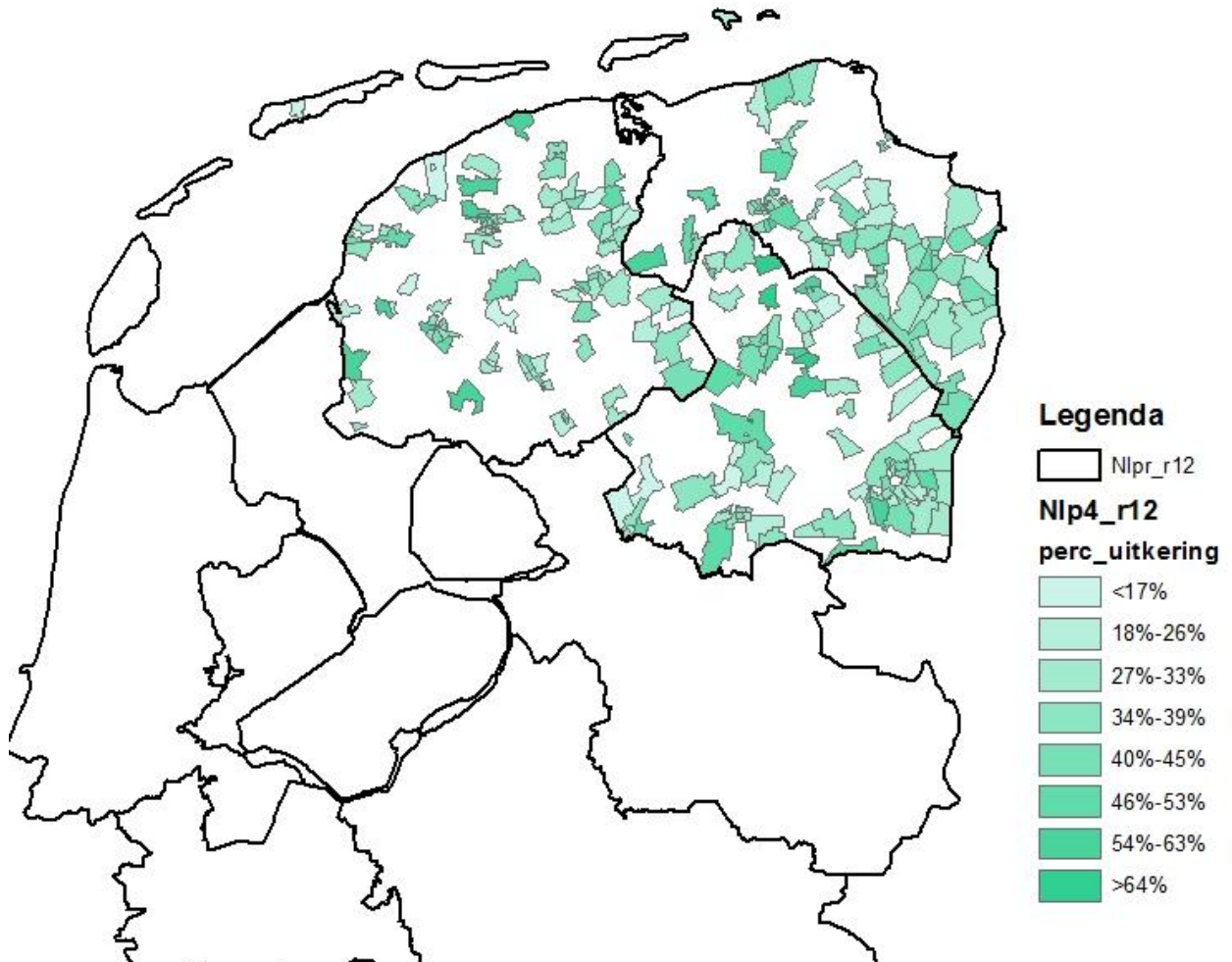
## 9.2 Bijlage II



Percentage niet-westerse allochtonen Noord-Nederland (bron: eigen bewerking)



### 9.3 Bijlage III



Percentage uitkeringontvangers Noord-Nederland (bron: eigen bewerking)